التحليل الإحصائي باستخدام

برنامج SPSS

د. مهدى محمد القصاص

كلية الآداب – جامعة المنصورة

2014 – 2013

الفصل الأول

مقدمة لبرنامج الحزم الإحصائية SPSS



(1-1) مقدمة:

علم الإحصاء أداة أساسية لا غنى عنها لتوصيف البيانات وتحليلها وحساب التقديرات والتنبؤات المستقبلية ونظراً لكبر حجم البيانات التي يتعامل معها علم الإحصاء من جهة واعتماده على أساليب كمية مطولة من جهة أخرى فإن استخدام الحاسب الآلي أصبح ضرورياً لإنجاز العمليات الإحصائية اختصارا للجهد والوقت.

ويعتبر البرنامج Statistical Package for Social Science (SPSS) أقدم البرنامج الإحصائية وأكثرها استخداماً من قبل شريحة واسعة من الطلبة والباحثين في مختلف التخصصات الإحصائية والطبية والهندسية والزراعية و الاجتماعية والتربوية والنفسية. ونظراً لقلة عدد من يجيد استخدام برنامج الـ SPSS بصورة وافية إضافة إلى افتقار المكتبة العربية إلى كتب تعليمية حول هذا البرنامج فقد كان هدفنا ومن خلال هذا الجزء أن نضع مصدراً تفصيلياً بين أيدي الباحثين في مختلف المجالات وطلاب مرحلة البكالوريوس في أقسام الإحصاء والتخصصات المختلفة الاخرى.

ويتضمن هذا الجزء تعريفاً بالجوانب الأساسية لحزمة SPSS ويهدف إلي إكساب العديد من المهارات اللازمة لتحقيق الاستفادة القصوى من إمكانيات البرنامج المتاحة آخذين في الاعتبار أن هناك عدد كبيراً من الدارسين ليس لديهم خلفية إحصائية وافية تمكنهم من التعامل مع البرنامج بصورة صحيحة.

كما يتضمن هذا الجزء نواحي ادارة الملفات و استخدام البرنامج كقاعدة بيانات فيما يتعلق بدمج الملفات وترتيبها واختيار الحالات وتبادل البيانات مع البرامج الأخرى كذلك الرسوم البيانية والجداول التكرارية والمقاييس الوصفية. كما يتضمن عرضاً موجزاً للجانب النظري للأسلوب الإحصائي المستخدم بالإضافة إلى التفسير الإحصائي لمخرجات البرنامج لبعض التطبيقات الإحصائية المهمة مثل اختبارات الفرضيات، الاختبارات المعلمية و اللامعلمية، تحليل التباين ، الانحدار...الخ. وقد تم التعامل مع تطبيقات البرنامج من خلال أمثلة مبسطة تتيح للقارئ الانتقال إلى خطوات متقدمة بسهولة ومعظم هذة الأمثلة مأخوذ من مصادر عربية وأجنبية معتمدة.

ويبحث علم الإحصاء في طرائق جمع البيانات وتحليلها وتفسيرها من خلال مجموعة من الطرائق الرياضية أو البيانية. وتهدف هذه العملية إلى وصف متغير أو مجموعة من المتغيرات من خلال مجموعة من البيانات(العينة) والتوصل بالتالي إلى

قرارات مناسبة تعمم على المجتمع الذي أخذت منه هذه العينة. ومن المعروف أن جمع المعلومات من جميع أفراد المجتمع أمر شاق يصعب تحقيقه في كثير من الأحيان، فذلك يحتاج إلى وقت وجهد ومال كثير، أما أخذ عينة عشوائية وممثلة من هذا المجتمع فعملية أسهل وتحتاج إلى جهد ووقت ومال اقل.

والبحث الذي يستخدم الأساليب الإحصائية للخروج بالنتائج والقرارات لا بد أن يمر في عدة خطوات.

أولا: تحديد المشكلة أو هدف الدراسة بوضوح ودقة، لأنه إذا كان هدف الدراسة غير واضح كانت النتائج غامضة وغير دقيقة.

ثانيا: تحديد الأداة التي ستستخدم لجمع البيانات وهي هنا الإستبانة.

ثالثا: تحديد العينة التي ستجمع منها البيانات وطرائق جمعها.

رابعا: ترميز البيانات (Coding) وتحويلها إلى أرقام أو حروف حتى يسهل إدخالها إلى الحاسوب ويسهل التعامل معها، ومن ثم إجراء التحليلات الإحصائية حسب التحليلات الإحصائية حسب أهداف البحث المنشود.

وقبل تناول عمليات الإدخال والتحليل لابد من مراجعة الركائز الأساسية لعلم الإحصاء (المتغيرات – اختيار العينة– تصميم الإستبانة)، لان هذه الركائز تحدد إلى حد كبير نوع التحليل الإحصائي المنشود.

أولا: طرق اختيار العينة من مجتمع

قبل أن نبدأ بكيفية اختيار عينة من مجتمع سنتعرف على الأسباب التي تجعلنا نختار عينة من مجتمع، بمعنى آخر **هناك عدة اعتبارات قد تستدعي استخدام أسلوب المعاينة، ومن بينها:**

1. تجانس المجتمع مثل المواد السائلة حيث لا يوجد ما يبرر أجراء فحص لكل أفراد المجتمع.
2. عوامل الوقت والجهد والتكلفة والملائمة بدون التضحية بدقة النتائج إلى حد كبير.
3. تعرض الوحدات المستخدمة في الاختبار للتلف عند فحص المجتمع كاملا (بيض، مصابيح الإضاءة، قوة مقاومة سيارة للمقاومة).

4. تعذر حصر أفراد المجتمع لأسباب عملية مثل فحص اتجاهات جميع المستهلكين حول سلع معينة أو توجهات الرأي العام حول قضايا عامة اقتصادية أو سياسية.

تعريف المجتمع: المجتمع هو مجموعة العناصر أو الأفراد التي ينصب عليهم الاهتمام في دراسة معينة وبمعنى آخر هو جميع العناصر التي تتعلق بها مشكلة البحث وقد يكون مجتمع الدراسة طلاب جامعة معينة أو سكان إقليم معين ، فمثلا إذا كانت مشكلة الدراسة هو ضعف توصيل المياه إلى المباني العالية (أكثر من ثلاث أدوار) في مدينة غزة فان مجتمع الدراسة أو البحث هو جميع المباني المرتفعة الأكثر من ثلاث أدوار في مدينة غزة، ويعتبر كل مبنى مؤلف من أكثر من ثلاثة أدوار مفردة البحث.

تعريف العينة: العينة هي مجموعة جزئية من المجتمع، ويكون حجم العينة هو عدد مفرداتها وعادة تجرى الدراسة على العينة.

أنواع البيانات الإحصائية:

كلما كان جمع البيانات دقيقا زادت ثقة الدارس في الاعتماد عليها، ولا يكون تحليل البيانات صحيحا أو مفيدا إذا كان هناك أخطاء في جمع البيانات، وهناك نوعين من البيانات وهما:

1- البيانات النوعية:

نحصل على هذا النوع من البيانات عندما تكون السمة (الخاصية) تحت الدراسة هي سمة نوعية والتي يمكن تصنيفها حسب أصناف أو أنواع وليس بقيم عددية مثل تصنيف الجنس إلى ذكر وأنثى، وتصنيف كليات الجامعة إلى طب وهندسة وعلوم وتجارة وآداب وتجارة وغيرها ، وتستخدم عدة مقاييس لقياس البيانات النوعية منها:

(أ) التدرج الاسمي

هذا المقياس يصنف عناصر الظاهرة التي تختلف في النوعية لا في الكمية، وكثيرا ما نستخدم الأعداد لتحديد هوية المفردات، وفي هذه الحالة لا يكون للعد ذلك المدلول الكمي الذي يفهم منه عادة. فمثلا يمكن استعمال العددين 0، 1 ليدلا على التصنيف حسب الجنس فيجعل الصفر يدل على الذكر و الـ 1 يدل على الأنثى، لاحظ أن 0، 1 لا يدلان على قيم عددية أي لا يخضعان للعمليات الحسابية لأنه يمكن تعيين أي عددين بدلهما ليدلا على نوع الجنس. وأمثلة أخرى على المقياس الاسمي : الحالة الاجتماعية

(أعزب– متزوج) ، ونوع العمل (إداري – أكاديمي – عمل آخر) . ويجدر بالذكر أن هذا المقياس لا يعطي الأفضلية لإحدى طبقات المجتمع على الأخرى.

(ب)التدرج الترتيبي

يقع هذا التدرج في مستوى أعلى من التدرج الاسمي، فبالإضافة إلى خواص التدرج الاسمي فان التدرج الترتيبي يسمح بالمفاضلة، أي بترتيب العناصر حسب سلم معين: مثل الرتب الأكاديمية (أستاذ (1)، استلذ مشارك(2)، أستاذ مساعد (3)، محاضر(4)، مدرس(5)، معيد(6)) وتقديرات الطلاب (ممتاز(5)، جيد جدا(4)، جيد(3)، مقبول(2)، راسب(1)) ، وكذلك درجة التأييد لإجابة السؤال (موافق بشدة (5)، موافق (4)، متردد(3)، لا أوافق (2)، لا أوافق بشدة (1))ويجدر بالذكر أن هذا المقياس لا يحدد الفرق بدقة بين قيم الأفراد المختلفة.

2- البيانات الكمية أو العددية

عندما تكون السمة تحت الدراسة قابلة للقياس على مقياس عددي فان البيانات التي نحصل عليها تتألف من مجموعة من الأعداد وتسمى بيانات كمية أو عددية، مثل علامات الطلاب في امتحان ما أو كميات السلع المستوردة، أجور العاملين في مصنع معين، وغيرها كثير.....

طرق جمع البيانات الإحصائية:

يتم جمع البيانات الإحصائية بإحدى الطرق التالية:

1- **طريقة المسح الشامل:** فيها تجمع البيانات من جميع مفردات المجتمع دون استبعاد أي مفردة، فمثلا إذا أردنا التعرف على مستوى طلاب الجامعة الإسلامية في مادة الإحصاء نقوم برصد درجات جميع طلاب القسم في مادة الإحصاء وهكذا...
وهذه الطريقة عادة تكون طويلة ومكلفة وتحتاج إلى الكثير من الوقت ناهيك عن عدم إمكانية تطبيقاتها في الحالات التي تؤدي فيها جمع البيانات عن مفردات البحث إلى فناء هذه المفردات.

2- **طريقة العينة:** وفيها يتم اختيار عينة تمثل المجتمع وتجرى عليها الدراسة وتعمم النتائج على المجتمع وكلما كانت العينة مختارة بطريقة صحيحة وممثلة تمثيلا صادقا المجتمع كلما كانت النتائج صادقة ودقيقة.

طرق اختيار العينة :

تصنف طرق المعاينة إلى الطرق غير العشوائية والطرق العشوائية أو الاحتمالية.

طرق اختيار العينة غير العشوائية

تكون العينات في هذه الطريقة انتقائية ولا تمثل المجتمع تمثيلا صحيحا، وإنما تتم وفق اختيار الباحث، ولذلك لا تكون هناك فرصة متساوية لأفراد المجتمع في الظهور في العينة، وهذه العينات تستخدم بهدف الحصول على نتائج استطلاعية نظرا لان اختيار عينات عشوائية يتطلب وقتا أو تكلفة أو جهود كبيرة.وفي هذه العينات لا يمكن استخدام أساليب الإحصاء التحليلي والذي يقتصر استخدامه على العينات العشوائية، ومن العينات الغير عشوائية ما يلي:

1. **العينات العرضيةAccidental samples** وتحدث عندما يتم جمع بيانات من المواطنين أو العمال في مصنع كبير الذين يصادفونهم حول اتجاهاتهم نحو سلع معينة أو نحو إدارة مصنع أو نظم الرقابية فيه للحصول على بعض المعلومات والمؤشرات بأقل تكلفة أو جهد ممكن.
2. **المعاينة الطبقية غير العشوائية Quota sampling** : وتحدث على سبيل المثال عندما يقسم مجتمع الدراسة في مصنع إلى طبقة الإداريين وطبقة العمال، أو إلى إناث وذكور، وبذلك تراعى نسبة المجموعات الفرعية في الدراسة. ولكن العينة من كل طبقة لا تأخذ بطريقة عشوائية وإنما يقوم الباحث باختيار الذين يصادفهم.
3. **العينة الغرضية Purposive sampling** : والتي تستخدم عند دراسة تكاليف صناعة على سبيل المثال، الأمر الذي يتطلب تعاونا من المستجوب لتوفير المعلومات.

طرق اختيار العينات العشوائية Random sampling

تسمح طرق اختيار العينات العشوائية بالحصول على عينات ممثلة للمجتمع، ويكون احتمال سحب أي مفردة معروفا ومتساويا ويمكن حسابه ولذلك تسمى عينة احتمالية فمثلا إذا كان حجم العينة المختارة 25 مفردة من مجتمع حجمه 500 فان احتمال سحب كل مفردة هو $\frac{25}{500} = 5\%$

تعريف العينة العشوائية: هي العينة التي يكون فيها احتمال اختيار جميع المفردات متساوي ومعروف ويمكن حسابه.

وهناك طرق مختلفة للاختيار العينة من أهمها:

1- العينة العشوائية البسيطة Sample random sampling

تتصف العينة العشوائية البسيطة بأنها مجموعة جزئية من المجتمع الأصلي وبحجم معين لها نفس الفرصة (الاحتمال) لتختار كعينة من ذلك المجتمع، ويمكن الحصول على عينات عشوائية بسيطة باستعمال جداول الأعداد العشوائية وسنوضح مثال اختيار عينة عشوائية باستخدام الجداول في المحاضرة.

2- العينة المنتظمة: Systematic sampling

يرى الكثيرون أن طريقة المعاينة المنتظمة هي في جوهرها شكل من أشكال المعاينة العشوائية البسيطة. وتعرف العينة المنتظمة بأنها العينة التي تأخذ بحيث يتم إضافة رقم معين بشكل منتظم من قائمة كاملة مرتبة عشوائيا لأفراد المجتمع. وتعتبر العينة المنتظمة بديلا عن العينة العشوائية البسيطة للأسباب التالية:

(أ)العينة المنتظمة أكثر سهولة في التنفيذ من العينة العشوائية البسيطة.

(ب)العينة العشوائية يستطيع شخص غير مدرب لتعينها.

مثال: إذا أردنا اختيار عينة حجمها $n=200$ من مجموعة من بطاقات التسجيل في إحدى الجامعات التي يسجل فيها $N = 3000$ طالبا لندرس البطاقات التي بها أخطاء.

الحل: إن طريقة العينة المنتظمة تقتضي بان يكون طول الفترة الذي سيسحب منها أول مفردة بطريقة عشوائية وهي $\frac{3000}{200}=15$. ولذلك نختار رقما عشوائيا من 1 إلى 15 وليكن 8.

نختار الرقم 8 ومن ثم نضيف 15 للرقم 8 وبذلك نسحب الرقم 23 ، ثم نضيف الرقم 15 للرقم 23 لنسحب الرقم 38، وهكذا .... وتكون آخر بطاقة مسحوبة هي رقم 2993.

ونلاحظ هنا انه إذا لم يكن طول الفترة عددا صحيحا فإننا نقرب الجواب إلي عدد صحيح.

3- العينة الطبقية العشوائية Stratified random sampling

تستخدم هذه الطريقة عندما يكون المجتمع منقسما إلى طبقات طبيعية وتكون لدينا الرغبة في تمثيل جميع هذه الطبقات في العينة. ونعرف العينة المنتظمة كالتالي:

تعريف العينة المنتظمة العشوائية: هي العينة التي تؤخذ من خلال تقسيم وحدات المجتمع إلى طبقات متجانسة واختيار عينة عشوائية بسيطة أو منتظمة من كل منها.

وتتلخص الطريقة بتحديد حجم العينات الجزئية المتناسبة من كل طبقة على أساس المعادلة

حجم العينة الطبقية = (حجم الطبقة ÷ حجم المجتمع) × حجم العينة .

مثال: إذا كانت طبقات أحد المجتمعات تحتوي العناصر كما في الجدول التالي:

الطبقة الأولى	الطبقة الثانية	الطبقة الثالثة	الطبقة الرابعة	الطبقة الخامسة
500	400	280	200	220

وأراد باحث اختيار عينة حجمها 150 من هذا المجتمع، فما حجم العينة في كل طبقة.

الحل: حجم المجتمع الكلي = 500 + 400+ 280+ 200+ 220 = 1600

حجم العينة من الطبقة الأولى = $500\times\frac{160}{1600} = 50$

حجم العينة من الطبقة الثانية = $400\times\frac{160}{1600} =40$

حجم العينة من الطبقة الثالثة = $280\times\frac{160}{1600} = 28$

حجم العينة من الطبقة الرابعة = $200\times\frac{160}{1600} = 20$

حجم العينة من الطبقة الخامسة = $220\times\frac{160}{1600} = 22$

ثانيا: جمع البيانات:

هناك عدة طرق لجمع البيانات نذكر منها:

1- المقابلة الشخصية Personal Interview

وهي أن تقوم بمقابلة أفراد العينة والتحدث إليهم عن الموضوع الذي يتم إجراء البحث فيه وبذلك فان كمية المعلومات التي سنقوم بجمعها ستكون دقيقة إلى حد ما، إلا أن تحليلها سيكون صعبا، وعليك أن تنتبه إلى تدوين البيانات أثناء المقابلة لان أي خطا في تدوين هذه البيانات يؤدي إلى خطا في النتائج.

2- الملاحظة المباشرة Direct Observation

عندما لا يكون هناك أفراد للعينة، فانك تستخدم هذه الطريقة أي الملاحظة المباشرة، ومن الأمثلة عليها أن تقف على تقاطع طرق، وتعد السيارات التي تمر من هذا التقاطع من الساعة الثامنة وحتى التاسعة بهدف حصر كثافة السير في وقت ذهاب الموظفين إلى أعمالهم، أو أن تقوم بمراقبة تصرف مجموعة من الأطفال أثناء اللعب وتدوين الملاحظات بهدف التعرف على سلوكيات الأطفال في بعض المواقف.

3- الإستبانة Questionnaire

الإستبانة هو وسيلة لجمع البيانات اللازمة للتحقق من فرضيات المشكلة قيد الدراسة، أو للإجابة على أسئلة البحث، وعند تصميم الإستبانة يجب مراعاة بعض الشروط حتى تضمن دقة النتائج وصحتها، ومن أهم هذه الشروط:

- يجب أن تكون أسئلة الإستبانة بسيطة ومفهومة للجميع بنفس الطريقة ولا تكون غامضة.

مثال: كم عدد الأطفال لديك ؟

هنا يتحير المجيب ليسال هل الطفل من هو دون سن الخامسة أم السابعة أم العاشرة...

ولذلك على الباحث أن يعيد السؤال ليصبح مثلا:

كم عدد الأطفال الذين تقل أعمارهم عن 12 سنة لديك..؟

- يجب على الباحث أن يبتعد عن تلك الأسئلة التي توحي بالإجابة. وغالبا ما تكون الأسئلة المنفية موحية بالإجابة

نعم ☐ لا ☐

مثال: ألا تعتقد أن أسلوب هذا الكتاب مبسط للدارس ؟

فالمجيب سيقوم باختيار الإجابة الأولى، وكان الباحث يريد أن يقوم المستجيب بالإجابة كما يريد الباحث.

- **يجب تحديد الكميات أو الوحدات عندما تكون الإجابات أرقاما.**

مثال: **كم تحتاج من كمية الماء للشرب يوميا؟ ....**

سيجيب أحد الأشخاص لتر ماء ويجيب آخر 5 كئوس ، أو ...

لذلك يعاد صياغة السؤال إلى كم لترا من الماء تشرب في اليوم؟ ...

- **يجب أن تكون الأسئلة مباشرة وواضحة وان لا يفكر المستجيب بعمق ليجيب على الأسئلة.**

- **يجب أن تكون الإستبانة قصيرة قدر الإمكان، حيث قد لا يكون عند المجيب وقتا طويلا لإجابة أسئلة الإستبانة.**

- **يفضل أن توزع الإستبانة على مجموعة صغيرة للتجريب وتعديل الأخطاء قبل التطبيق النهائي.**

- **يجب أن تكون الإستبانة صادقة وثابتة، فان لم تكن صادقة فلن تكون المعلومات دقيقة. أما إذا لم تكن الإستبانة ثابتة فلن نستطيع تعميم الإستبانة، ولن يكون قرارنا صالحا لفترة من الزمن وسنوضح كيفية التأكد من صدق أسئلة الإستبانة ودرجة ثباتها من خلال برنامج SPSS.**

ثالثا: الترميز (عملية الانتقال من الاستبيان إلى برنامج SPSS)

الخطوة التالية والتي تسبق إدخالها إلى الحاسوب بهدف التحليل هي ترميز البيانات. وترميز البيانات هي عملية تحويل إجابات كل سؤال إلى أرقام أو حروف يسهل إدخالها إلى الحاسوب.

حسب مفهوم SPSS فان الأشخاص (المشاهدات) الذين يقومون بالإجابة على أسئلة الاستبيان يطلق عليهم اسم حالات (Cases) ، وكل سؤال (فقرة) في الاستبيان هو عبارة عن متغير (Variable) ، وتسمى إجابات الأشخاص على الأسئلة (الفقرات) بقيم المتغيرات (Values of Variables).

يحتوي الاستبيان على عدة أنواع من الأسئلة، وهذه الأنواع هي:

أ) سؤال يسمح باختيار إجابة واحدة فقط:

نعم ☐ لا ☐

مثال: هل أنت مواطن أم لاجئ ؟

متغير واحد يكفي لتمثيل هذا السؤال، في هذه الحالة نرمز للإجابة " نعم " بالرمز 1 وللإجابة " لا " بالرمز 2 أو نرمز للإجابة" نعم " بالرمز N وللإجابة" لا " بالرمز Y ولكن يفضل استخدام الترميز الأرقام لان عملية إدخال البيانات الرقمية في SPSS تتم بسهولة اكثر ولان الحاسوب يفرق بين الحروف الصغيرة والكبيرة وكذلك فالمر فان كثير من الأوامر في SPSS تنفذ فقط مع المتغيرات الرقمية ولا تنفذ مع المتغيرات الحرفية.

مثال: **هل توافق أن يكون تسجيل الطالب في الجامعة عبر الحاسوب؟**

موافق بشدة		موافق		محايد		معارض		معارض بشدة	

في هذا المثال ربما يستخدم الرقم 5 ليدل على الإجابة " موافق بشدة" والرقم 4 ليدل على الإجابة " موافق" والرقم 3 ليدل على الإجابة " محايد" والرقم 2 ليدل على الإجابة " معارض" والرقم 1 ليدل على الإجابة " معارض بشدة".

ب) سؤال يسمح بأكثر من إجابة:

مثال: ما هي أهم الهوايات التي تمارسها ؟

القراءة		الرياضة		السباحة		الصيد		غير ذلك	

في هذا السؤال نلاحظ أن الشخص يمكن أن يعطي اكثر من إجابة، لذلك فان متغيرا واحدا لا يكفي لتمثيل السؤال. في هذه الحالة يفضل إنشاء خمسة متغيرات، كل متغير له احتمال إجابتين نعم / لا ويستخدم لهما 1 للإجابة " نعم " و 0 للإجابة " لا"

مثال: رتب القنوات الفضائية التالية حسب أهميتها لك.

الجزيرة		المنار		الفلسطينية		العربية		الكويتية		السورية	

في هذا السؤال يجب إنشاء ستة متغيرات وإعطاء الرقم 6 للقناة الأكثر أهمية والرقم 5 للأقل أهمية إلى أن نصل إلى اقل القنوات أهمية وإعطائها الرقم 1.

ج) سؤال مفتوح جزئياً:

ويقصد بذلك السؤال الذي يسمح للشخص باختيار إجابة موجودة ضمن الخيارات أو كتابة إجابة أخرى غير موجودة ضمن الخيارات.

مثال: عند سفرك للخارج أي خطوط الطيران تستخدم؟

الفلسطينية		المصرية		القطرية		الأردنية		غير ذلك اذكرها	.....

في هذا النوع من الأسئلة فان متغيرا واحدا يكفي لتمثيل هذا السؤال لان المسموح به هو إجابة واحدة فقط(شريطة أن يستخدم المسافر شركة طيران واحدة) إلا أن عملية تعيين

رموز تصف قيم المتغير (الإجابات) هي صعبة نوعا ما وتتم باستخدام عدة **طرق يمكن تلخيصها كالتالي:**

الطريقة الأولى: أن ترمز لكل شركة طيران وردت بالإجابة برقم من 1 إلى N حيث يمثل N عدد شركات الطيران الواردة بالإجابة وهذه طريقة سيئة لأنها تحتاج لوقت كبير، لأنه سيتعامل مع كل استبيان بشكل منفرد ليتم جمع البيانات كلها.

الطريقة الثانية: تعيين الرمز 5 ليصف الإجابة " غير ذلك " بحيث يتم معاملة هذه الإجابات كمجموعة واحدة عند تحليل الإجابات بغض النظر عما ذكر من أنواع شركات الطيران الممكنة. وهذه الطريقة سيئة لأنها تمكننا من فقدان معلومات كثيرة، إلا أن هذا الفقدان من المعلومات قد لا يكون مشكلة إذا كان الاستبيان يركز على شركات الطيران الواردة في السؤال.

ولاختيار أي الطرق أفضل فإنه يجب الأخذ بعين الاعتبار العوامل التالية :

- الهدف من الإستبانة
- شكل الاستبيان الذي تم تقديمه للأشخاص وكيفية الإجابة علية.
- الوقت المتاح للباحث.
- الدعم المادي المتوفر للباحث.
- الدقة المطلوبة.

(2-1) الدخول للبرنامج (Getting Start):

يوجد لدينا طريقتين للدخول إلى البرنامج:

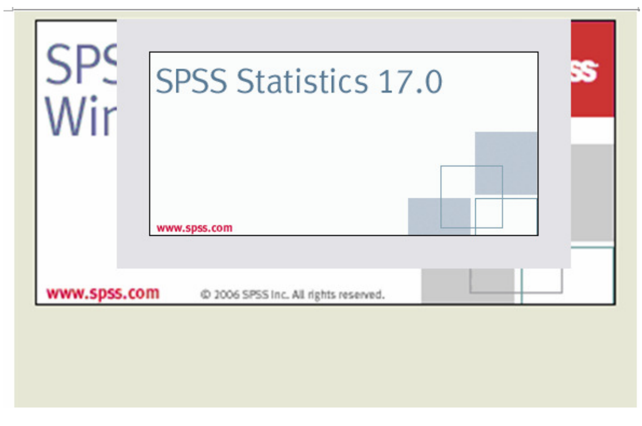
الطريقة الأولى:

بعد تحميل البرنـامج ننقر علـى Start فـي شريط المهام ومن ثم النقر على All Programs ثم اختيار SPSS كما في الشكل (1-1) التالي.



الشكل(1-1)

نبدأ الدخول إلى البرنامج وذلك بظهور الشاشة الافتتاحية للبرنـامج وهو إعـلان عن الحزمة لمدة ثوان، كما في الشكل (2-1) :

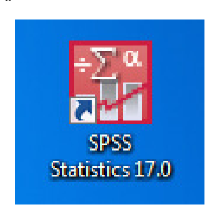


الشكل (2-1)

بعد ظهور الشاشة الافتتاحية تبدأ أول شاشة من شاشات الحزمة في الظهور شكل (1-4)

الطريقة الثانية:

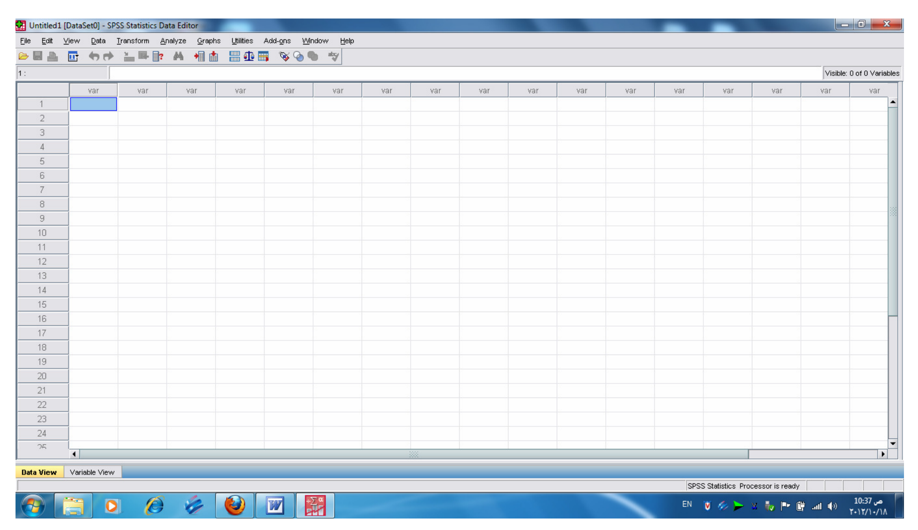
الدخول بهذه الطريقة يتم عن طريق أيقونة مختصرة (Short Cut Icon) أعدت ووضعت على سطح المكتب عن طريق المستخدم ، وبالنقر مرتين عليها فندخل للبرنامج مباشرة. وهذه الأيقونة تكون كما بالشكل الأتي:



الشكل (1-3)

(1-3) نافذة محرر البيانات (Data Editor):

تعرض هذه النافذة محتويات ملف معين من البيانات حيث يمكن انشاء ملف جديد أو تحرير ملف موجود وإن هذه النافذة تفتح تلقائيا عند بدء تشغيل البرنامج كما بالشكل (1-4) التالي:



الفصل الثاني

تجهيز ملفات البيانات

Data File Preparation

(2-1) مقدمة:

في هذا الفصل نناقش طريقة تعريف المتغيرات وإدخال البيانات محل الدراسة للبرنامج تمهيداً لإجراء التحليل الإحصائي عليها.

(2-2) المتغير والحالة (Variable and Case):

المثال التالي يعطي فكرة مبسطة عن المتغير والحاله.

مثال (2-1):

البيانات التالية تمثل قيد مجموعة معينة من الأشخاص في اختبار معين.

جدول (2-1)

Name	ID	Gender	Age	Birth date	Income
Mohammed	1	1	38	1/1/1970	1000
Mervat	2	2	27	20/8/1980	500
Ahmad	3	1	48	15/3/1960	2000
Heba	4	2	22	15/4/1985	300
Nabil	5	1	18	1/1/1990	200
Noha	6	2	58	20/7/1949	3000

جدول بيانات التحليل: كل عمود في المثال يمثل متغير من المتغيرات، والمتغيرات يمكن تقسيمها كما يلي:

- المتغير الأول: متغير الاسم ID وهو متغير رمزي.
- المتغير الثاني: متغير الجنس Gender (الرقم 1 يمثل الذكور والرقم 2 يمثل الإناث).
- المتغير الثالث: متغير العمر Age وهو متغير رمزي أيضا.
- المتغير الرابع: متغير تاريخ الميلاد Birth date وهو متغير تاريخ .
- المتغير الخامس: متغير مستوي الدخل Income وهو متغير رقمي.

- **أنواع البيانات في البرنامج:**
  - **بيانات كمية**: يستعمل للبيانات العددية (القابلة للقياس الكمي) مثل بيان الدخل ودراجاتك في الامتحان وهكذا.

- **بيانات مكودة**: يستعمل لقياس المتغيرات الاسمية والتي لا يمكن ترتيبها تصاعديا أو تنازليا كما لا يمكن اجراء العمليات الحسابية على هذا النوع من المتغيرات مثل / تقسيم المجتمع إلى ذكور وإناث و الأسماء أيضا يمكننا اعتبارها مقياس اسمي.
- **بيانات تاريخ**: وهي البيانات التي نرغب بتوضيح بدايتها ، نهايتها، أو فترتها بالتاريخ الزمني.
- **بيانات عملات**: وهي البيانات التي نرغب بتوضيح العملات المستخدمة في حسابها.
- البيانات النصية: وهي البيانات التي نرغب في كتابتها بشكل نصي في البرنامج مثل اسم المتغيرات.

- **تعريف الاكواد** (Codes):

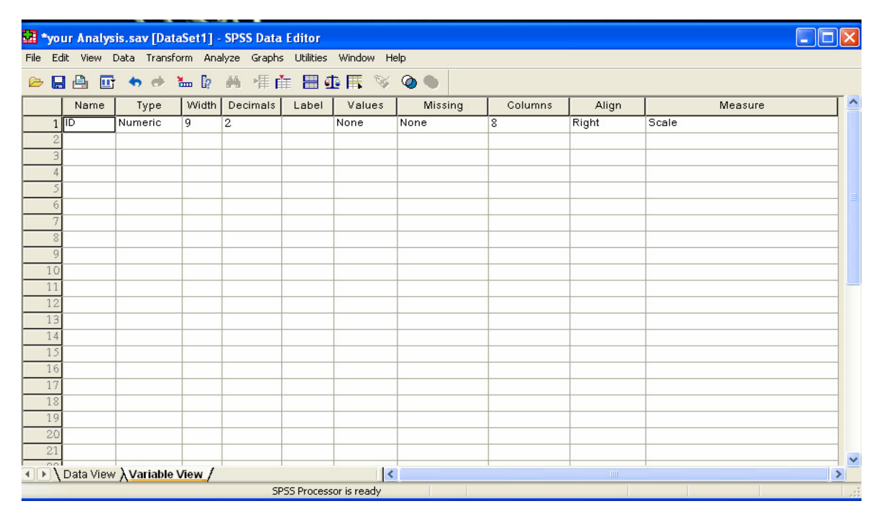
نقصد بالأكواد تحويل المتغير اللفظي إلى رقمي ويطلق على الرقم الذي يمثل قيمة المتغير اللفظي كود أو رمز. يمكن التمييز بين المتغيرات سواء كانت لفظية أو رقمية بناء على ما يسمى وحدة القياس .

(3-2) تعريف المتغيرات (Defining Variables):

تعريف المتغير في برنامج SPSS يعنى أنه يجب تحديد ما يلي:

- اختيار اسم المتغير Variables Names.
- اختيار مميز المتغير Variable label.
- اختيار قيمة المميز إن وجد Value label.
- تعريف القيم المفقودة Missing Values.
- تحديد نوع المتغير Variable Type.
- تحديد شكل عرض البيانات في الأعمدة Column format .

حيث نقوم بتعريف المتغير من شاشة Variable View الظاهرة في الشكل التالي:



شكل (2-1)

وسنقوم بشرح كل عمود من أعمدة الشاشة السابقة لكون تلك الشاشة شاشة تعريف متغيرات تحليلك الإحصائي.

- **اختيار اسم المتغير:**

تبدأ عملية تعريف المتغير بتحديد اسما للمتغير في العمود الأول من أعمدة Variable View والظاهر في الشكل وذلك بالنقر عليه نقراً مزدوجاً بزر الماوس الأيمن وكتابة اسم المتغير بعد ذلك باستخدام لوحة المفاتيح الخاصة بجهازك.

توجد عدة شروط يجب أن تؤخذ في الاعتبار عند تحديد أسم المتغير هي:

- ✓ يجب أن لا يزيد عن 64 حرفاً (for version 15 or 12) أو ثمانية حروف أو أرقام لنسخ السابقة.
- ✓ أن يبدأ الاسم بحرف (ليس رقماً) ولا يكون به فراغات .
- ✓ لا يسمح باستخدام الحروف التالية كجزء من اسم المتغير (. / ! / ؟ / ، / *).
- ✓ لابد أن يكون الاسم وحيد في نفس الملف بمعنى أن كل متغير يأخذ اسم لا يأخذه متغيراً آخر.
- ✓ لا يمكن استخدام الكلمات الآتية كاسم للمتغير With – Ge – Not – Ne – Eq – To – Le -And – Gt – Or - It by – all لأن هذه الكلمات لها استخدامات معينة داخل الحزمة SPSS .

✓ يمكن كتابة الاسم باللغة الإنجليزية سواء بحروف كبيرة أو صغيرة.

- **تعريف نوع المتغير** (Kind of Variable):

هذا هو العنصر الثاني من تعريف المتغيرات محل اهتمامك في التحليل الإحصائي إذ وهو نوع المتغير وهو العمود الثاني من Variable View وفي الحقيقة يعتبر هذا العنصر في منتهى الأهمية بالنسبة للتحاليل الإحصائية حيث تختلف أنواع المقاييس الإحصائية المستخدمة على حسب نوع المتغيرات محل الدارسة. ومعظم الاختيارات تكون رقمي "numeric" هذا يعني أن المتغير يأخذ رقماً لاحظ الآتي أن كلمة رقماً لا تعني بتأكيد قيمة عددية يمكننا من خلالها إجراء العمليات الحسابية المعتادة ونجد أن هناك خيار آخر وهو "string" وهو خاص بالمتغيرات التي تأخذ نصوص في التحليل الإحصائي والجدول الآتي يوضح أنواع البيانات المتاحة في برنامج SPSS:

جدول (2-2)

نوع المتغير	مثال
Numeric	1000.36
Comma	000.005،1
Scientific	1*e3
Dollar	000.00،$1
String	Mervat

ونلاحظ أن البرنامج يضع بشكل تلقائي نوع المتغير Numeric ويمكننا تغيرها للأسباب الآتية:

- أننا نرغب في توضيح فواصل عشرية كثيرة أو قليلة جداً.
- أن أرقام المتغير كبيرة جداً مثل 12456789123 فهنا على سبيل المثال نستطيع اختيار الخيار comma لكي نستطيع أن نقرأ الرقم فيصبح الرقم 123،789،456،12، أو نضعها باستخدام scientific فتصبح 12.45*E9 فنستطيع قراءتها بسرعة على أنها 12 بليون (E3 هي ترمز إلي الآلاف، E6 ترمز إلى المليون ، و E9 ترمز إلى البليون).

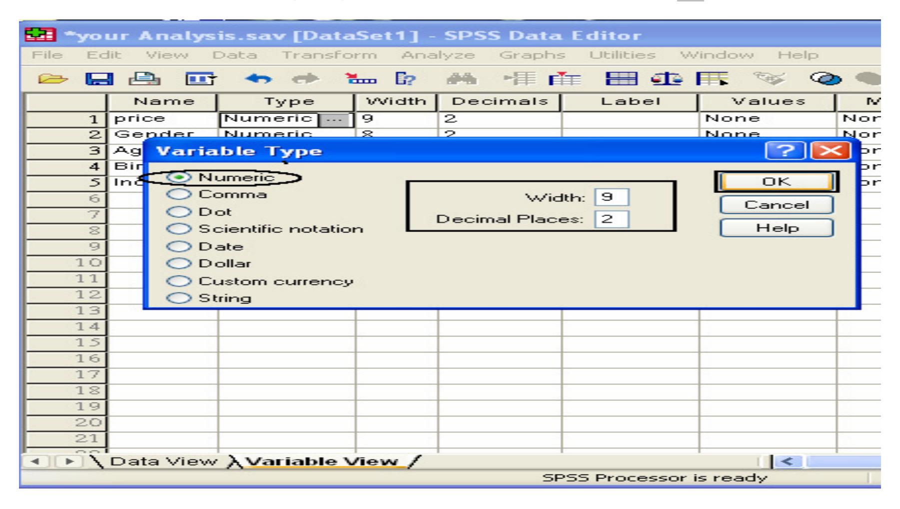
⌖ في حالة ما تكون المتغير هي عملات ونرغب في توضيح ذلك فإننا نضع نوع المتغير Dollar.

والآن نستعرض بعض الأمثلة لتوضح كيف يمكن تغير نوع المتغير في البرنامج:

مثال (2-1) بيانات كمية:

لتغير نوع البيانات إلى رقمي نتبع الخطوات الآتية:

نختار Variable view ⇦ وبعد ذلك عند المتغير الذي نريد أن نغير نوعه نختار العمود الثاني عنده والذي يحمل عنوان Type ونلاحظ أن بجوار نوع المتغير توجد نقاط على شكل " ... " نقوم بنقر عليها فيظهر الشكل (2-2) التالي:



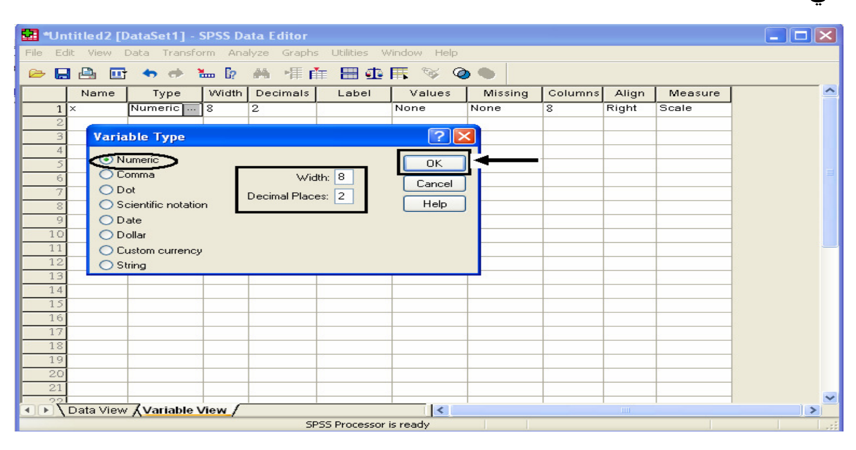
شكل (2-2)

نختار Numeric وهي أول خيار ونستطيع من ذلك عرض عمود البيانات لهذا المتغير بحيث يظهر عدد من خانات الأرقام أكبر وهنا نختار 9 Width وعدد الخانات العشرية 2 بحيث إذا كان المتغير يأخذ الرقم الآتي 123456789.659 فيظهر في البرنامج بالشكل الآتي 123456789.66 فيقوم البرنامج بتقريب الخانات العشرية إلي أقرب رقمين فقط . وبعد ذلك ننقر على Ok لتأكيد الاختيار ونلاحظ أننا بعد النقر على Ok يتغير قيمة العمود الثالث الذي يحمل عنوان Width إلى 9 وهو عدد خانات الرقم الأساسية، ويتغير أيضا قيمة العمود 4 والذي يحمل عنوان Decimals إلى 2 وهو عدد الخانات العشرية المتاحة .

مثال (2-2): البيانات المكودة.

البيانات المكودة هي البيانات الاسمية التي نعطي لها رقماً يعبر عنها في البرنامج وهي من أهم أنواع البيانات في البرنامج من هذه البيانات بيان النوع أو الجنس بيان الحالة الزوجية ففي بيان الجنس نرغب بتعبير عن حالة النوع" ذكر" برقم" 1" وحالة النوع أنثى برقم" 2" ونلاحظ أن تلك الأرقام هي أكواد وليست قيماً يمكننا إجراء العمليات الحسابية عليها. ولتعريف البرنامج بالبيانات المكود نتبع الخطوات التالية:

نختار Variable view ⇦ وبعد ذلك عند المتغير الذي نريد أن نغير نوعه نختار العمود الثاني عنده والذي يحمل عنوان Type نقوم بنقر عليها فيظهر الشكل (2-3) التالي:



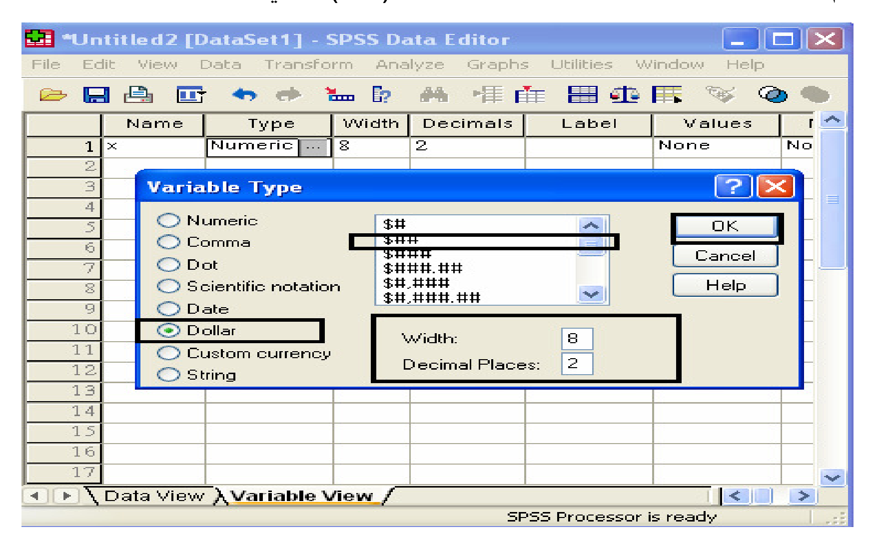
شكل (2-3)

نختار Numeric وهي أول خيار ونستطيع من ذلك عرض عمود البيانات لهذا المتغير بحيث يظهر عدد من خانات الأرقام أكبر وهنا نختار Width 2 وعدد الخانات العشرية 0 حيث أننا لدينا خانتين فقط وليس لدنيا أرقام عشرية وبعد ذلك النقر على Ok لتأكيد الاختيار.

مثال (3-2): بيانات العملات

في حالة البيانات التي تعبر عن عملات معينة نريد توضيح أن ذلك البيان هو بيان يتبع عملة معينة مثل بيان الدخل فإننا نعرف ذلك البيان إلى البرنامج على طريق أختيار Variable view ⇦ وبعد ذلك عند المتغير الذي نريد أن نغير نوعه نختار العمود الثاني عنده والذي يحمل عنوان Type نقوم بنقر عليها ومن ثم نقوم بتحديد

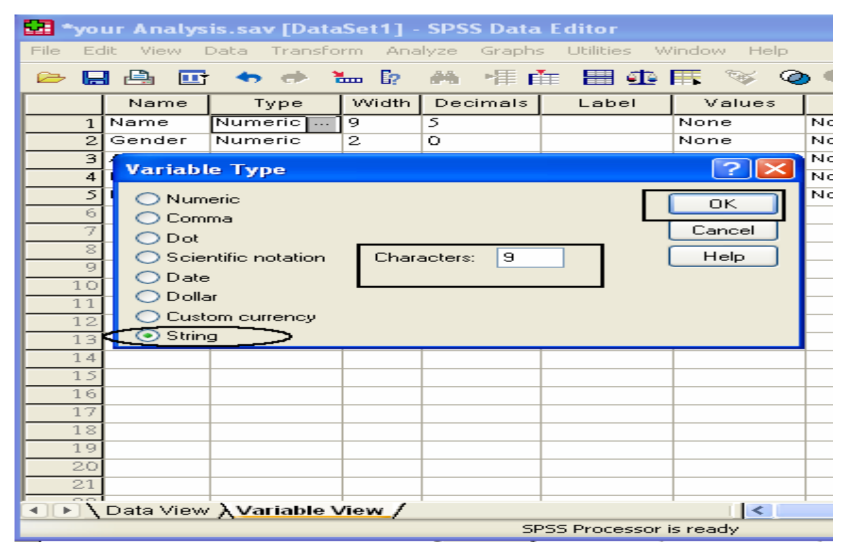
الخيار Dollar يلي ذلك تحديد عدد خانات بيان العملة المرغوب توضيحها وبعد ذلك نقوم بنقر على ok لتأكيد الاختيار فيظهر الشكل (4-2) التالي:



شكل (4-2)

مثال (4-2): بيانات نصية

في حالة البيانات النصية إذا أردنا كتابة المتغيرات النصية في البرنامج بشكل نصي فإننا نعرف ذلك البيان إلى البرنامج على طريق الخطوات التالية: نختار Variable view ⇦ وبعد ذلك عند المتغير الذي نريد أن نغير نوعه نختار العمود الثاني عنده والذي يحمل عنوان Type نقوم بنقر عليها ومن ثم نقوم بتحديد الخيار string يلي ذلك تحديد عدد خانات البيان النصي المرغوب وضعه وهو أقصى عدد أحرف يمكن إضافتها وبعد ذلك نقوم بنقر على ok لتأكيد الاختيار فيظهر الشكل (5-2) التالي:

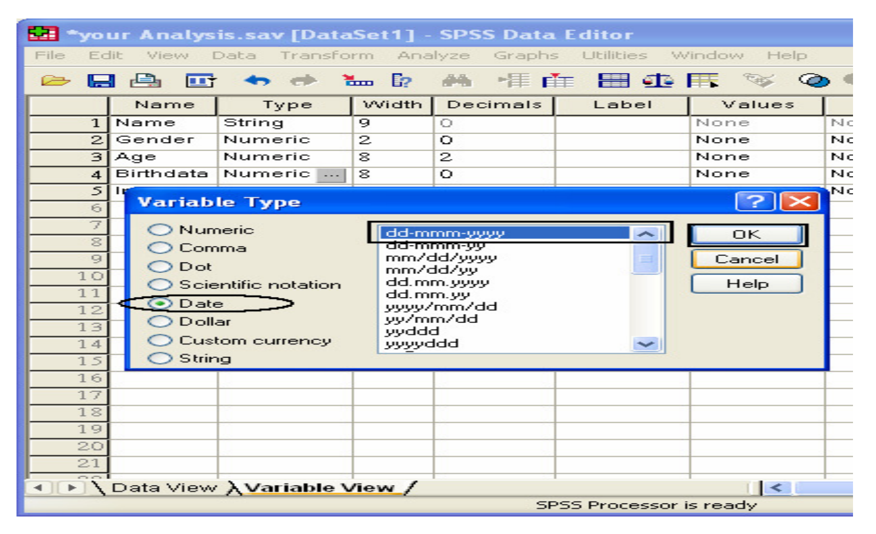


شكل (2-5)

مثال (2-5): بيانات التاريخ

إذا أردنا كتابة متغير بيان تـاريخ الميلاد في البرنـامج فإننا نعرف ذلك البيان إلـى البرنامج على طريق الخطوات التالية:

نختار Variable view ⇦ وبعد ذلك عند المتغير الذي نريد أن نغير نوعه نختار العمود الثاني عنده والذي يحمل عنوان Type نقوم بـالنقر عليها ومن ثم نقوم بتحديد الخيار data يلي ذلك تحديد نمط التاريخ، فإذا افترضنا أننا نتبع النظام البريطاني في كتابة التـاريخ حيث يكتب اليوم في البدايـة ومن ثم كتابـة الشـهور وفـي النهايـة كتابـة السنوات بكامل الشكل فإننا ننقر على النمط الأول لتاريخ وبعد ذلك نقوم بنقر علـى ok لتأكيد الاختيار فيظهر الشكل (2-6) التالي:

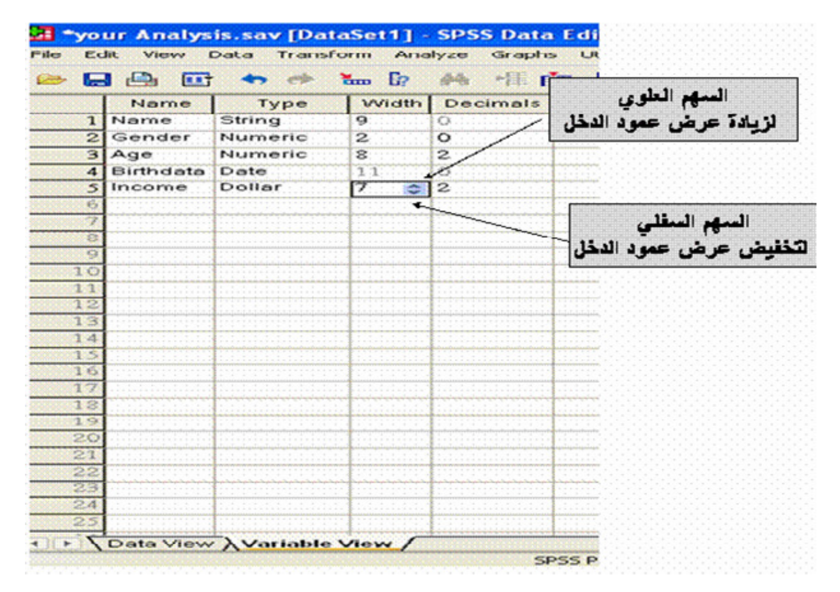


شكل (2-6)

تعريف (2-1)

عرض المتغير (Width Variable):

هـذا هـو العنصـر الثالـث مـن تعريـف المتغيـرات محـل اهتمامـك فـي التحليـل الإحصـائي إذ وهو عرض المتغير وهو العمود الثالث من Variable View وفـي الحقيقة يمكننا من تغير عرض المتغير مباشرة عندما نقوم بتعريف نوع المتغير أو بنقر مباشـرة علـى الخلية الموجـودة فـي عمـود عـرض المتغير وصـف المتغير المراد تغيـر عرضـه فيظهر سـهم إلـى أعلـى وسـهم إلـي أسفل وعنـد النقر علـى السـهم العلـوي يقـوم بزيـادة عـرض المتغير مقدار وحدة فـي كل مرة ضـغط وهكذا بالنسبة للأسـهم السـفلي حيث يقوم بتخفيض عرض المتغير بمقدار وحدة واحدة في كل مرة ضـغط والشكل (2-7) يوضح هذه العملية:

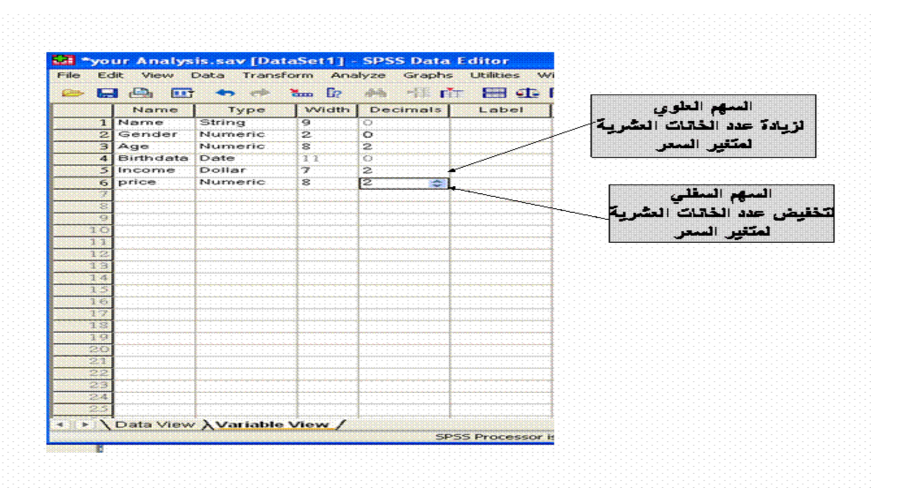


شكل (2-7)

تعريف (2-2)

عدد الخانات العشرية للمتغير (Decimals of Variable):

هذا هو العنصر الرابع من تعريف المتغيرات محل اهتمامك في التحليل الإحصائي إذ وهو عدد الخانات العشرية للمتغير وهو العمود الرابع من Variable View حيث يمكننا من تغيير عدد الخانات العشرية للمتغير مباشرة عندما نقوم بتعريف نوع المتغير أو بنقر مباشرة على الخلية الموجودة في عمود عدد الخانات العشرية للمتغير(Decimals) وصف المتغير المراد تغيير عدد خانته العشرية فيظهر سهم إلى أعلى وسهم إلى أسفل وعند النقر على السهم العلوي يقوم بزيادة عدد الخانات العشرية للمتغير بمقدار وحدة في كل مرة ضغط وهكذا بالنسبة للسهم السفلي حيث يقوم بتخفيض عدد الخانات العشرية له أيضا بمقدار وحدة واحدة في كل مرة ضغط والشكل (2-8) يوضح هذة العملية:

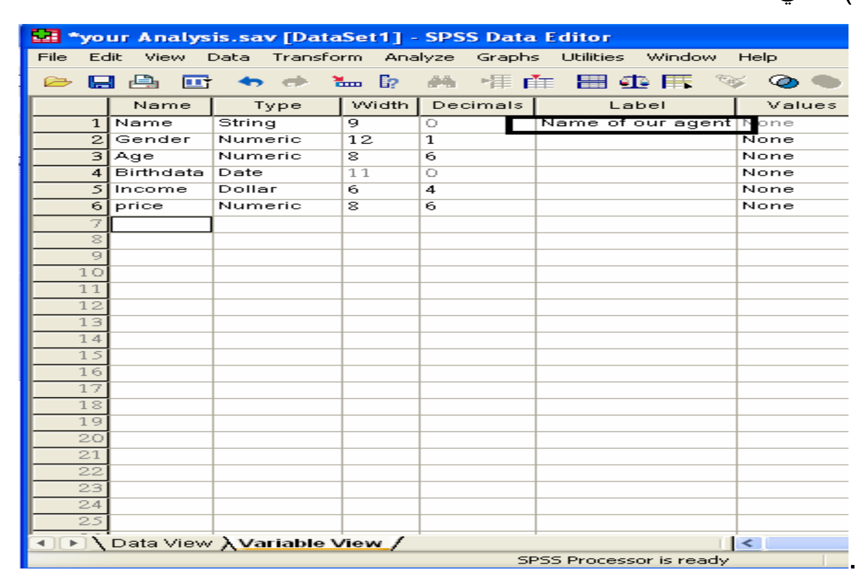


شكل (2-8)

تعريف (2-3)

مميز المتغير (Variable label):

يمكن أن يأخذ المتغير عنوانا يصل عدد رموزه إلى 256 رمز يستعمل لوصف المتغير مع إمكانية وضع فواصل ورموز خاصة بعكس تماماً خانة اسم المتغير فيمكنا من تعريف متغير الاسم على أنه اسم العميل Name of the agent كما بالشكل (2-9) التالي:



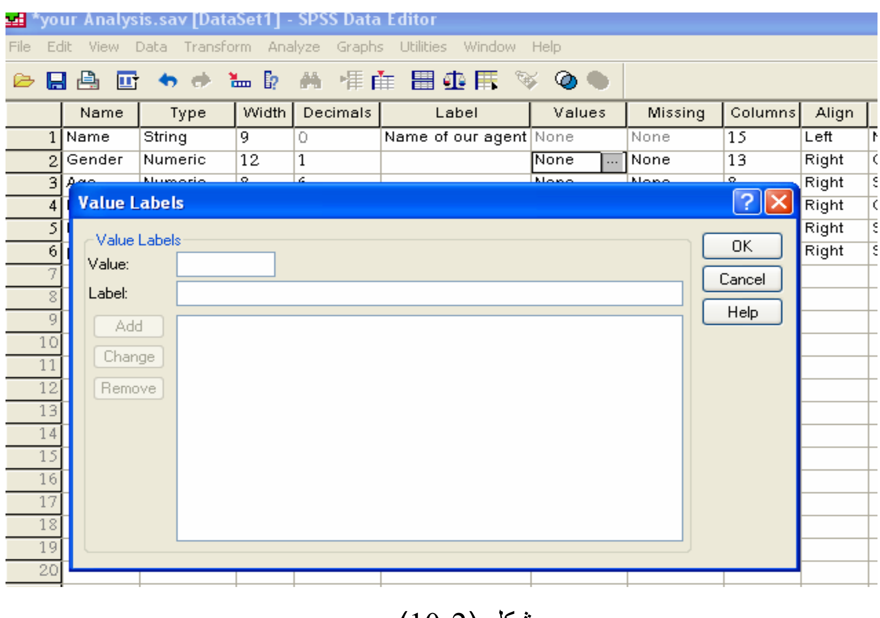
شكل (2-9)

تعريف (4-2)

قيمة المميز (Value label) :

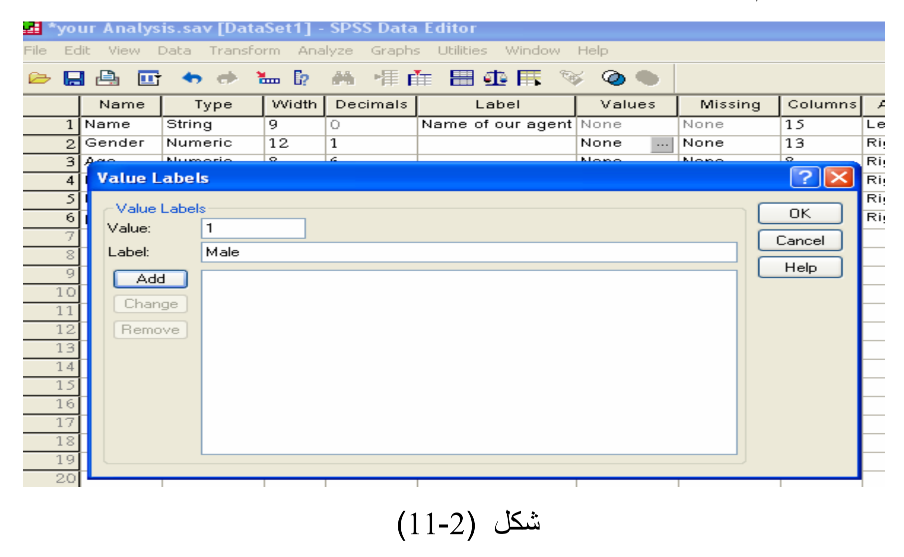
أحيانا تبرز الحاجة إلى تعيين عنوان للقيمة كون المتغير يستعمل قيما عددية للتعبير عن قيم غير عددية مثلا / متغير Gender يستعمل الرقم 1 للتعبير عن الذكور والرقم 2 للتعبير عن الإناث ونلاحظ أن هذا التعريف معرف لجميع المتغيرات ماعدا المتغيرات النصية ولذلك فإن في حالة المتغير النصي نجد أن قيمة المميز لا تكون منشطة تأخذ من خلال المثال رقم 4 أما باقي المتغيرات فيمكن اعطاء قيمة للتمييز لها، فعلى سبيل المثال متغير النوع Gender يمكننا من تعريف الكود"1" إلى الذكر والكود"2" للأنثى وذلك بإتباع الخطوات التالية:

نختار Variable view ⇦ وبعد ذلك عند المتغير الذي نريد أن نعطي له قيمة مميز له وليكن متغير النوع فنقف على الخلية الموجودة في العمود الذي يحمل عنوان "Value" وصف متغير النوع فنجده أن الخلية مكتوب بها فنقوم بنقر على النقاط التي تحمل اللون الرمادي فيظهر مربع الحوار الآتي:



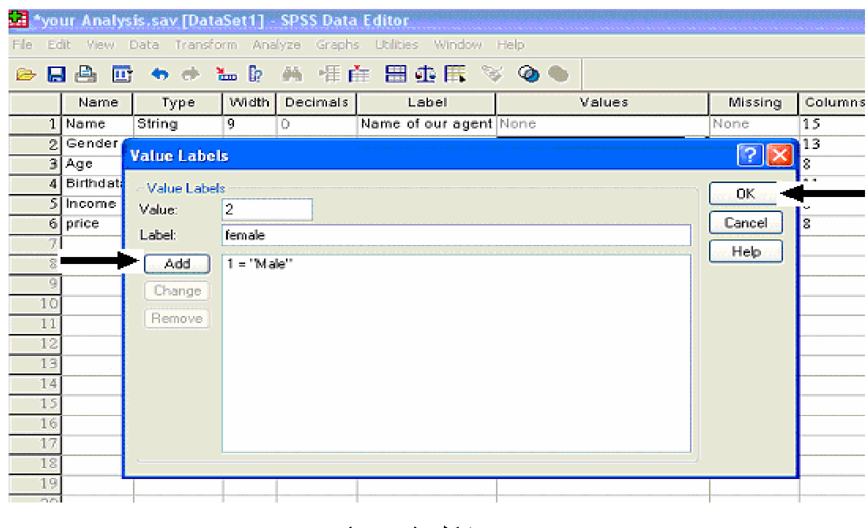
شكل (2-10)

فعندما نرغب في إضافة الكود "1" للذكر و2 للأنثى فإننا نكتب أمام الخانة Value القيمة "1" ونكتب أمام الخانة Label ذكراً أو Male فتظهر الشاشة التالية:

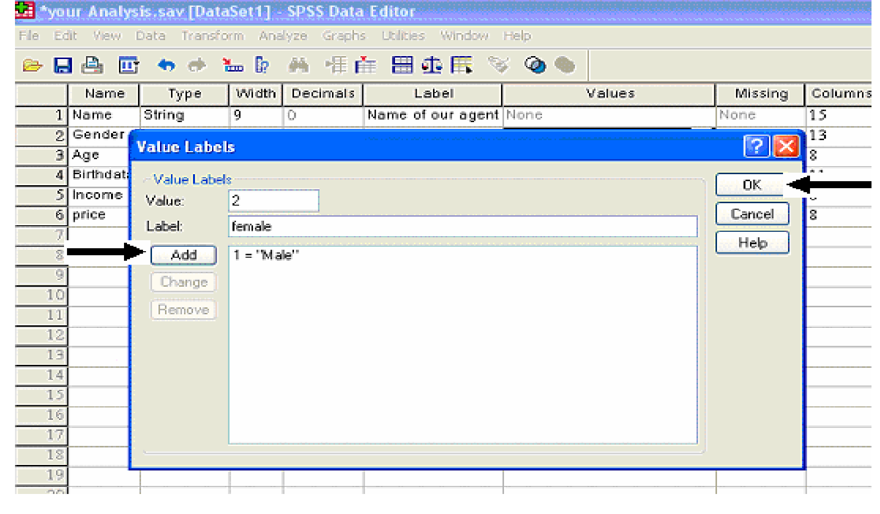


شكل (2-11)

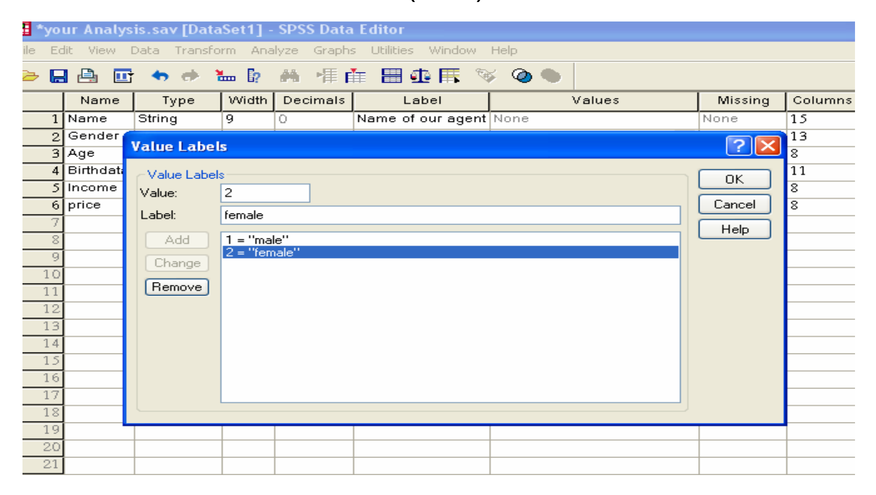
وبعد ذلك ننقر Add وذلك لإضافة الكود الخاص بالأنثى فعندما ننقر Add تظهر الشاشة فنكرر الخطوة السابقة كما فعلنا بالنسبة للكود الذكر ولكننا هنا في خانة Value نضع القيمة '2' وفي خانة Label نضع Female وفي حالة الانتهاء ننقر على Ok ونلاحظ أننا يمكننا حذف أي كود لا نرغب به فقط ننقر على الكود المراد حذفه وبعدها النقر على Remove الأشكال الآتية توضح ذلك.



شكل (2-12)



شكل (2-13)



ملاحظات:

- يمكن أن يصل طول عنوان القيمة إلى 60 رمزاً.
- يمكن أن يكون العنوان قيما عددية ليمثل قيما غير عددية مثلا الرقم 1 عنوان للذكور والرقم 2 عنوان للإناث.
- يمكن إظهار عنوان القيمة view variable أما بالنقر على الأيقونة في شريط الأدوات أو بتأشير Value label من قائمة view في شريط القوائم.

⌖ يظهر عنوان القيمة بدلاً من القيمة نفسها في مخرجات (جدول) برنامج SPSS ويعتبر هذا من أهمية توضيح عنوان للقيم في المتغيرات الاسمية أو الترتيبية.

تعريف (2-5)

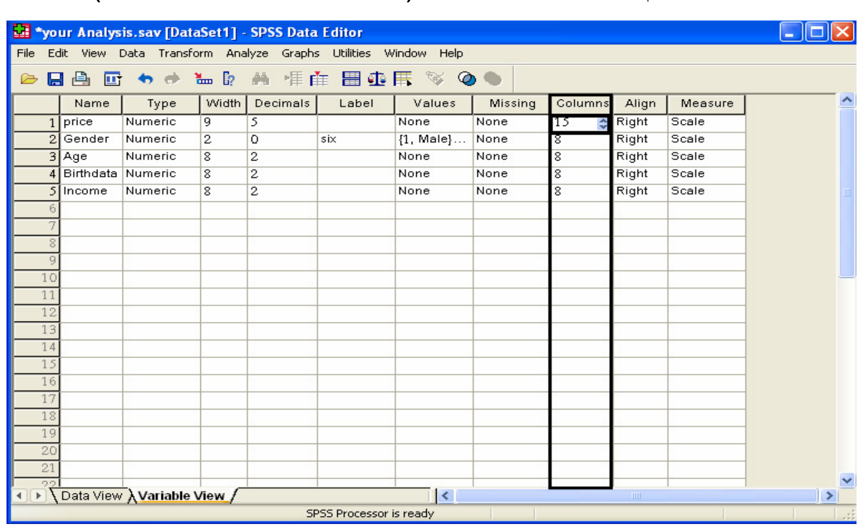
القيم المفقودة (Missing Values):

القيم المفقودة هي قيم موجودة أصلا، ولكننا لا نرغب إدخالها في التحليل الإحصائي كونها مثلاً قيماً شاذة أو أن نوع السؤال لا ينطبق على المستجيب.

تعريف (2-6)

عرض العمود (Column width):

يمكن تحديد عرض العمود لمتغير معين بالوقوف على الخلية الواقعة ضمن العمود المعنون Column في ورقة view variable حيث يمكن زيادة أو تقليل عرض العمود بواسطة الأسهم إلى الأعلى أو الأسفل (أو كتابة عرض العمود مباشرة).



شكل (2-14)

ملاحظات:

⌖ عرض العمود Column يمثل عدد الرموز المخصصة للعمود ويجب أن يكون عرض العمود أكبر أو يساوي عرض المتغير المضمن فيه.

⌖ يمكن تغير عرض العمود لمتغير معين في ورقة Data view مباشرة عن طريق نقر وسحب حدود العمود Clicking and dragging.

(4-2) محاذاة النص (Alignment):

لضبط محاذاة النص داخل خلايا المتغير انقر الخلية التابعة لمتغير معين في ورقة variable view الواقعة ضمن العمود المعنون Align ثم انقر السهم المتجه للأسفل لاختيار أمر مما يلي:

Left: لمحاذاة النص إلى يسار الخلية.

Center: لمحاذاة النص إلى وسط الخلية.

Right: لمحاذاة النص إلى يمين الخلية.

علماً أن المحاذاة الافتراضية هي (Right).

(5-2) أنواع وحدات القياس (Measures Types):

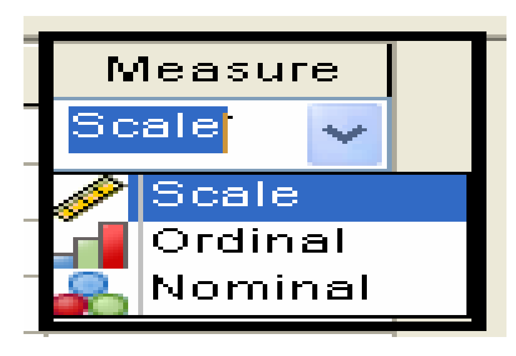
في البداية يجب تعريف أنواع وحدات القياس للبيانات قبل الشروع في تعريف ذلك لبرنامجك الإحصائي.

أنواع وحدات القياس:

- Nominal : يستعمل لقياس المتغيرات الاسمية وهي متغيرات لها عدد من الفئات دون أفضلية لإحداها على الأخرى ولذلك لا يمكن ترتيبها تصاعديا أو تنازليا كما لا يمكن اجراء العمليات الحسابية على هذا النوع من المتغيرات مثل / تقسيم المجتمع إلى ذكور وإناث.
- Ordinal : يستخدم لقياس المتغيرات الترتيبية وحيث أن هذا المتغير ذو عدد محدد من الفئات يمكن ترتيبها تصاعديا أو تنازلياً ولكن لا يمكن تحديد الفروق بينها بدقة مثلا تقدير طالب في الامتحان قد يكون" ممتاز ، جيد جدا ، جيد ، ... الخ" ويمكن أن يكون المتغير رمزيا أو عددياً على أنه يفضل الأخير (عددي).
- Scale : ويستخدم لقياس المتغيرات الكمية في فترة أو نسبة كدخل والعمر و الأسعار ... الخ.

ولغرض تعريف مقياس لمتغير معين انقر خلية المتغير التي تقع ضمن عمود measure في ورقة variable view حيث يظهر ثلاث خيارات scale، ordinal ،

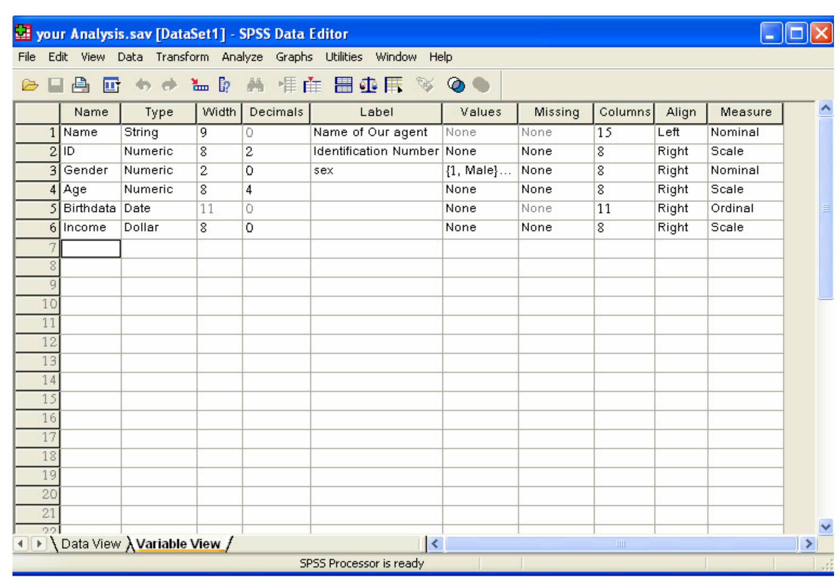
and nominal ونختار نوع المتغير على حسب التعاريف السابقة ومن ثم تظهر شاشة variable view كما يلي:



شكل (2-15)

وفي النهاية تظهر شاشة variable view.

نريد الآن ادخال بيانات الجدول (2-1) الموجودة في مثال (2-1) في بداية هذا الفصل. بعد تعريف المتغيرات:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Name	String	9	0	Name of Our agent	None	None	15	Left	Nominal
2	ID	Numeric	8	2	Identification Number	None	None	8	Right	Scale
3	Gender	Numeric	2	0	sex	{1, Male}...	None	8	Right	Nominal
4	Age	Numeric	8	4		None	None	8	Right	Scale
5	Birthdata	Date	11	0		None	None	11	Right	Ordinal
6	Income	Dollar	8	0		None	None	8	Right	Scale

شكل (2-16)

(2-6) إدخال البيانات (Entering Data):

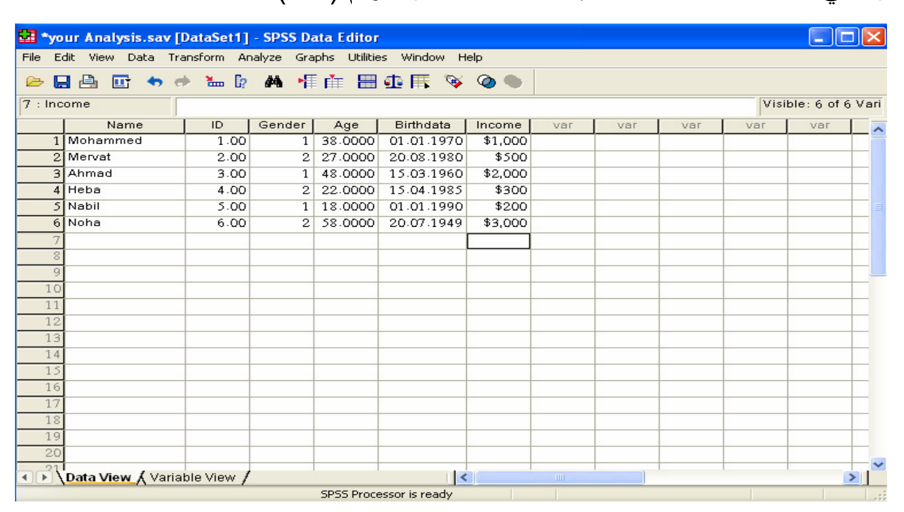
الآن وصلنا إلى مرحلة إدخال البيانات استعداداً لإجراء التحليل الإحصائي.

يتم إدخال البيانات كما يلي :

- نختار الخلية المراد إدخال قيمة المتغير فيها.
- نكتب القيمة المراد إدخالها في الخلية في مكان اسمه محرر الخلية Cell Editor موجود تحت شريط الأدوات.
- عند الانتقال للخلية التالية بالأسهم أو النقر على الأمر Enter تكتب القيمة المدخلة تلقائياً.
- عند حدوث أية أخطاء في عملية الإدخال يمكن إجراء التصحيح كالآتي :
- ننتقل إلى الخلية المراد تصحيحها ويتم النقر عليها.
- نكتب القيمة الصحيحة في محرر الخلية.

ملاحظة : ممنوع إدخال تعبيرات حسابية (جمع / ضرب / ...).

والآتي شكل data view بعد إدخال بيانات جدول رقم (2-1) فيه:



	Name	ID	Gender	Age	Birthdata	Income	var	var	var	var	var
1	Mohammed	1.00	1	38.0000	01.01.1970	$1,000					
2	Mervat	2.00	2	27.0000	20.08.1980	$500					
3	Ahmad	3.00	1	48.0000	15.03.1960	$2,000					
4	Heba	4.00	2	22.0000	15.04.1985	$300					
5	Nabil	5.00	1	18.0000	01.01.1990	$200					
6	Noha	6.00	2	58.0000	20.07.1949	$3,000					

شكل (2-17)

(7-2) عمليات مختلفة على البيانات:

- **تحديد المتغيرات والحالات:**

⌖ لاختيار (تحديد أو تظليل) متغير ما variable انقر الخلية التي تحتوي على اسم المتغير في أعلى العمود في ورقة Data view بزر الماوس الأيسر.

⌖ لاختيار حالة case بأكلمها انقر الخلية الحاوية على رقم الحالة في ورقة Data view بزر الماوس الأيسر.

⌖ لاختيار مجموعة من المتغيرات المتجاورة: انقر الخلية الحاوية اسم المتغير الأول.

⌖ النقر على مفتاح shift.

⌖ انقر الخلية الحاوية على اسم المتغير الأخير مع استمرار النقر على مفتاح shift.

⌖ لاختيار مجموعة من المتغيرات المتباعدة انقر بزر الماوس الأيسر الخلية الحاوية اسم المتغير الأول لاختياره.

⌖ انقر على مفتاح Ctrl مع استمرار النقر عليه (انقر الخلية الحاوية اسم المتغير الثاني لاختيارها وهكذا بالنسبة لباقي المتغيرات التالية).

⌖ بنفس الطريقة المستخدمة لاختيار مجموعة من المتغيرات المتجاورة أو غير متجاورة يمكن استعمالها لاختيار مجموعة من الحالات Cases المتجاورة أو غير المتجاورة.

- **إدخال (حشر) حالة جديدة** (Insert a New Case):

نستطيع حشر حاله بين حالتين موجودتين كالتالي:

⌖ ننتقل إلى (الصف) أسفل المكان المراد إضافة الحالة فيه ويتم النقر.

⌖ من القائمة Data ننقر على الأمر Insert Case فتظهرحالة جديدة خالية.

⌖ ندخل البيانات المراد إدخالها إلى الخلايا عن طريق محرر الخلية.

- **إدخال (حشر) متغير جديد** (Insert new variable):

الطريقة الأولى:

من شريط القوائم اختر القائمة Data ثم نختار Insert Variable فيضاف متغير جديد إلى يسار المتغير الحالي حيث يمكن تغيير الاسم .

الطريقة الثانية:

الوقوف على المتغير المراد إضافته على يسار المتغير وثم النقر بزر الماوس الأيمن واختيار insert variable وتسمية المتغير بعد ذلك في شاشة variable view بنفس الطريقة السابقة.

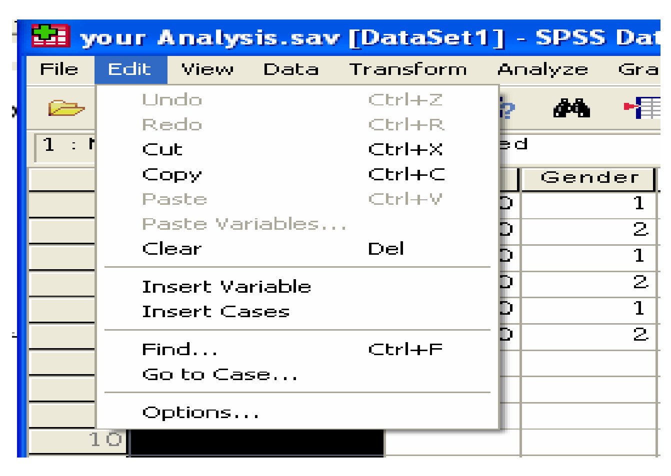
- **حذف صف** (Delete Row):

انقر اسم الحالة في ورقة Data View بزر الماوس الأيسر.

من القائمة Edit انقر على Clear تتم الإزالة الفورية للحالة. ويمكن استخدام الأمر Delete بعد اختيار المتغير مباشرة .

- **حذف عمود** (Delete Coulmn):

يتم حذف المتغير كالآتي حدد المتغير المراد حذفه من القائمة Edit التي تتضمن القائمة الآتية:



شكل (2-18)

انقر على Clear ويمكن استخدام الأمر Delete .

- **حذف صف** (Delete Case):

بنفس طريقة حذف المتغيرات يمكننا حذف الحالات الغير مرغوب فيها كالأتي:

انقر اسم الحالة في ورقة Data View بزر الماوس الأيسر.

من القائمة Edit انقر على Clear تتم الإزالة الفورية للحالة . ويمكن استخدام الأمر Delete بعد اختيار المتغير مباشرة .

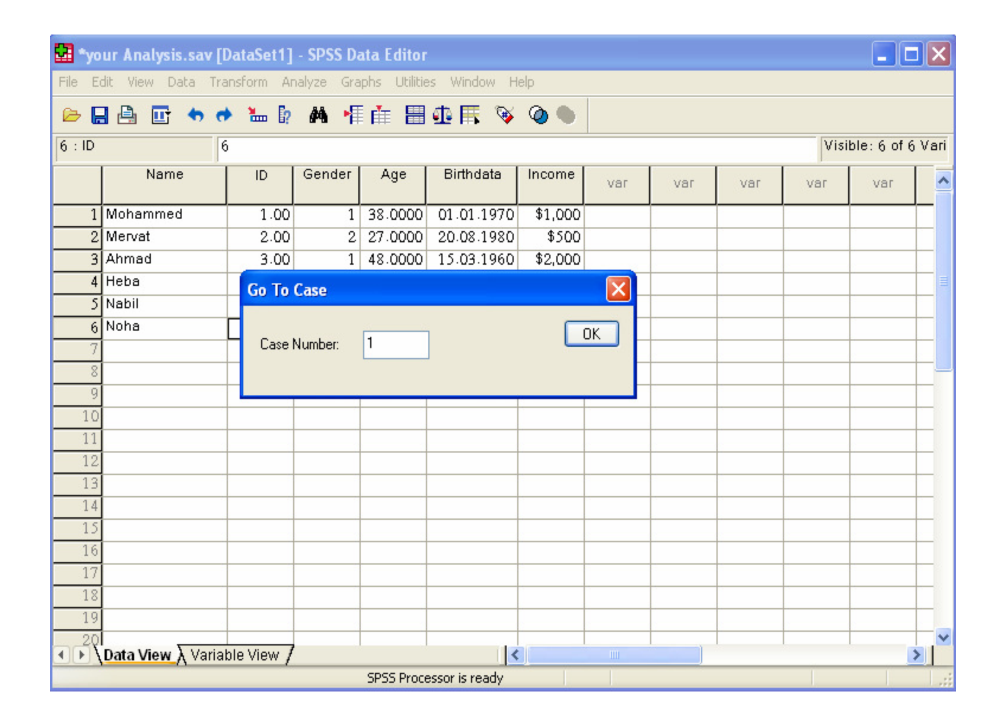
- **نسخ وقص ولصق الحالات والمتغيرات** (Paste of cases ، cut، copy and variable)

لعمل نسخة من متغير معين Copy نتبع الخطوات التالية:

- ⌖ حدد اسم المتغير ثم من شريط القوائم نختار edit ومن ثم اختيار copy من القائمة المدرجة انظر شكل (2-19) ويمكن النقر مباشرة على ctrl+c لإجراء عملية النسخة بشكل سريع نحدد بعدها المتغير المراد النسخ فيه.
- ⌖ من شريط القوائم نختار edit ومن ثم اختيار Past ويمكن النقر مباشرة على ctrl+v لإجراء عملية اللصق بشكل سريع. وعملية القص واللصق تتم بنفس الطريقة ، كما أن عملية نسخ ولصق وقص الحالات أيضا تتم بنفس الطريقة مع تبديل اختيار العمود (المتغير) بالصف (الحالة).

- **البحث عن حالات** (Go To Case):

- ⌖ انقر الأمر Go To Case من القائمة Data فيظهر المربع الحواري (انظر الشكل (2- 19) التالي:



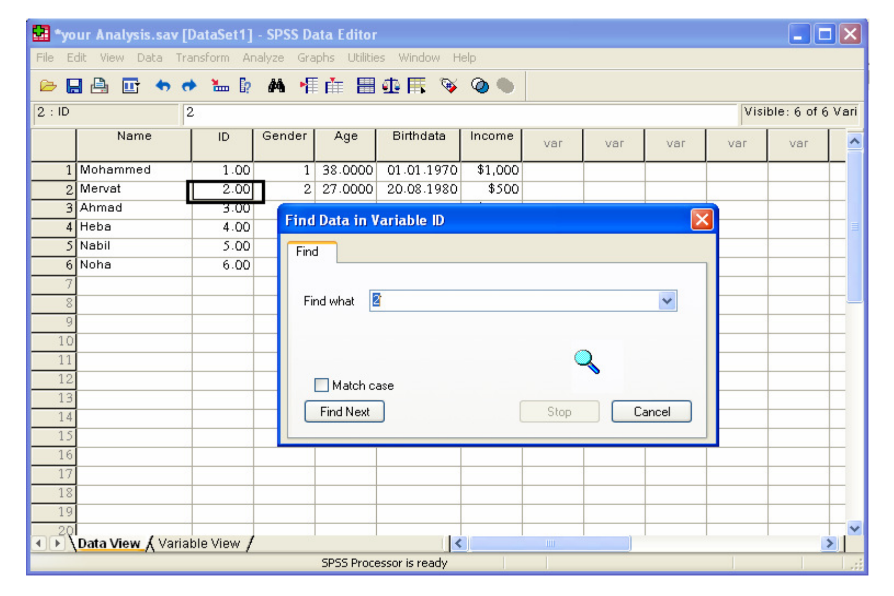
الشكل (2-19)

- في المربع Case Number اكتب رقم الحالة التي تريد الانتقال إليها.
- انقر Ok فيتم الانتقال إلى الحالة المحددة.

- **البحث عن قيمة** (Finding Values Finding Values):

تتم كالآتي:

- ننتقل إلى قمة العمود الخاص بالمتغير.
- انقر الأمر Find من القائمة Edit فيظهر المربع الحواري Search for Data الآتي (أنظر الشكل 2-20):



الشكل (2-20)

- في المربع الحواري حدد الرقم المطلوب البحث عنه في المربع Find What ونلاحظ أننا إذا أردنا البحث عن القيمة" 2 " فإننا نكتب القيمة" 2 " في خانة Find what وبعدها النقر على Find next فنلاحظ أن البرنامج يقف على أول حالة أو متغير لدية القيمة" 2 "ويحددها كما في الشكل السابق.

الفصل الثالث

معالجــة البيانات

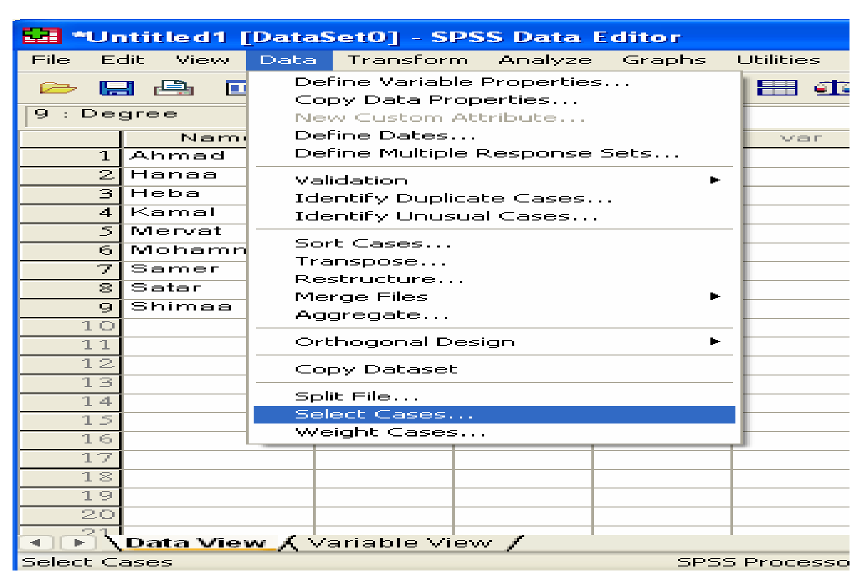
Data Processing

(1-3) مقدمة:

استعرضنا في الفصل السابق طريقة تعريف المتغيرات وإدخال البيانات محل الدراسة لبرنامج SPSS تمهيداً لإجراء التحليل الإحصائي عليها. في هذا الفصل نناقش بعض طريق معالجة البيانات وإجراء التحليل الإحصائي عليها.

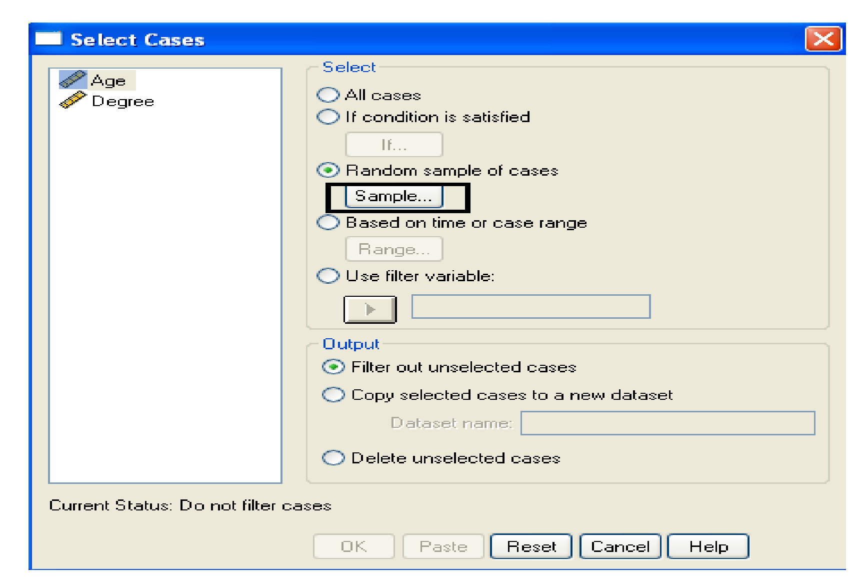
(2-3) اختزال (تخفيض) حجم البيانات Reducing sample size :

نفرض أننا لدينا 2 مليون بيان تم الحصول عليها من المشاهدات و لتخفيض حجم هذة البيانات، نقوم بإختيار عينة عشوائية منها لإجراء التحليل الإحصائي عليها. نفرض أنه تم اختيار 100 الف مشاهدة بطريقة العينة العشوائية البسيطة ولضمان عدم وجود تحيز في الاختيار نقوم باختيار Select Cases من القائمة Data كما بالشكل التالي:



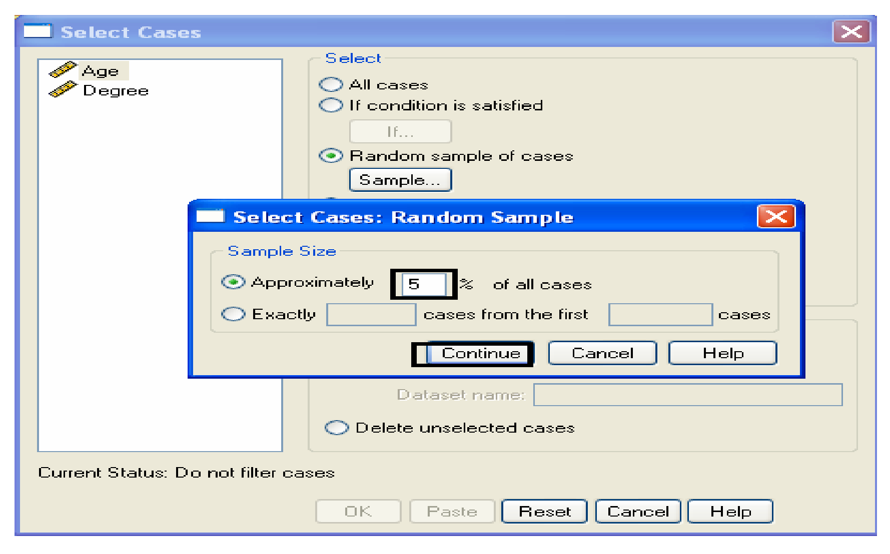
شكل (3-1)

بعدها تظهر الشاشة التالية :



شكل (3-2)

فنقوم باختيار Random sample of cases والنقر بعدها على مربع الحوار Sample و لأننا نختار 100 الف من 2 مليون بيان فإننا نحدد حجم العينة بتقريب بنسبة 5% من البيانات ومن ثم تظهر الشاشة التالية:



شكل (3-3)

(3-3) ترتيب البيانات Sort Data :

نقصد بعملية ترتيب البيانات بأن نرتب بيانات أحد المتغيرات (أو أكثر من متغير) في الملف ثم نرتب بقية المتغيرات تبعا لذلك. لترتيب بيانات ملف نذكر بعض الملاحظات الهامة :

- أن ترتيب بيانات الملف تتم بالنسبة لمتغير أو أكثر وليست مطلقة.
- أن عملية الترتيب تعني نقل الحالات الموجودة في الملف إلى أعلى أو إلى أسفل .
- لا يجوز نقل قيمة معينة (أو قيم) من حالة من مكانها إلى مكان آخر بل يجب نقل الحالة كلها.

مثال (3-1):

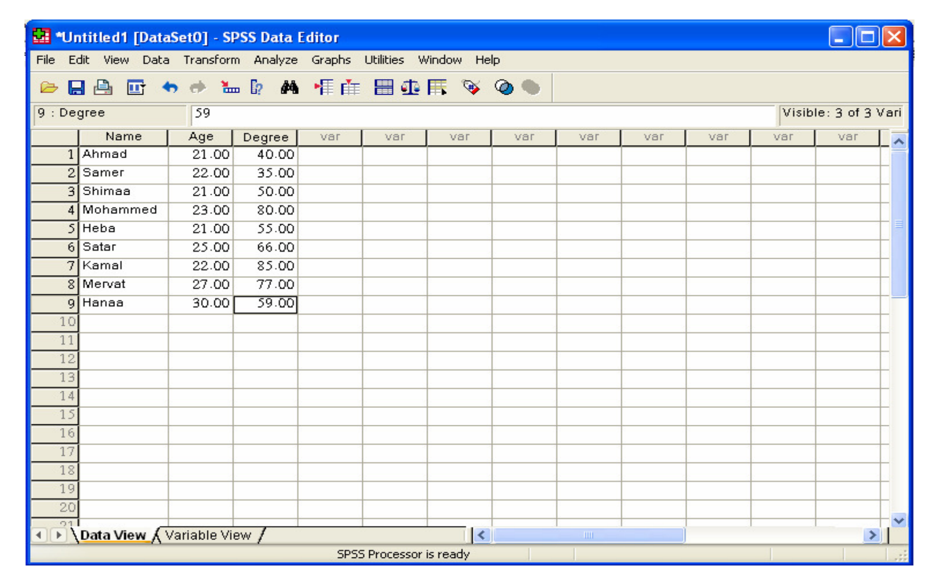
لدينا بيانات الملف التالي لخرجين أحد كليات العلوم:

جدول (3-1)

Name	Age	Total degree
Ahmad	21	40
Samer	22	35
Shimaa	21	50
Ali	23	80
Heba	21	55
Satar	25	66
Kamal	22	85
Mervat	27	77
Hanaa	30	59

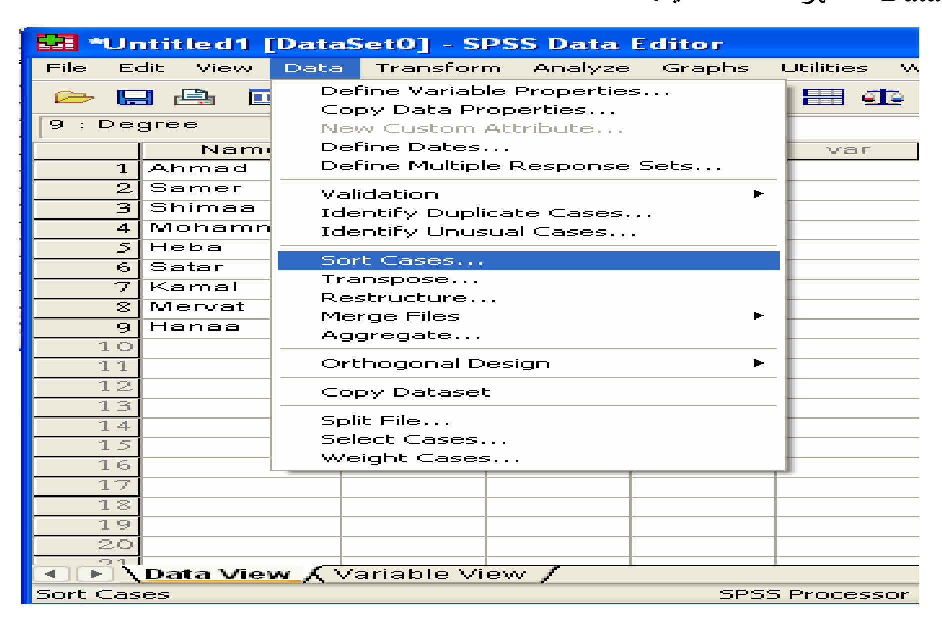
لتنفيذ ترتيب البيانات باستخدام البرنامج نتبع الخطوات التالية :

نقوم بإدخال البيانات في البرنامج فتظهر الشاشة التالية:



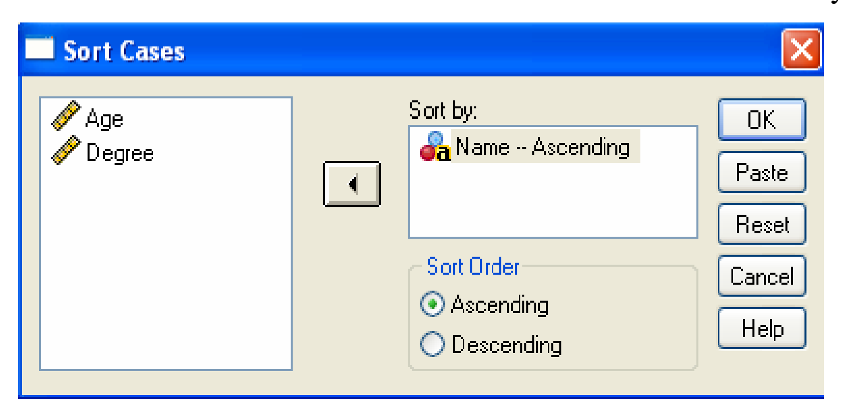
شكل (3-4)

انقر Data فتظهر الشاشة التالية:



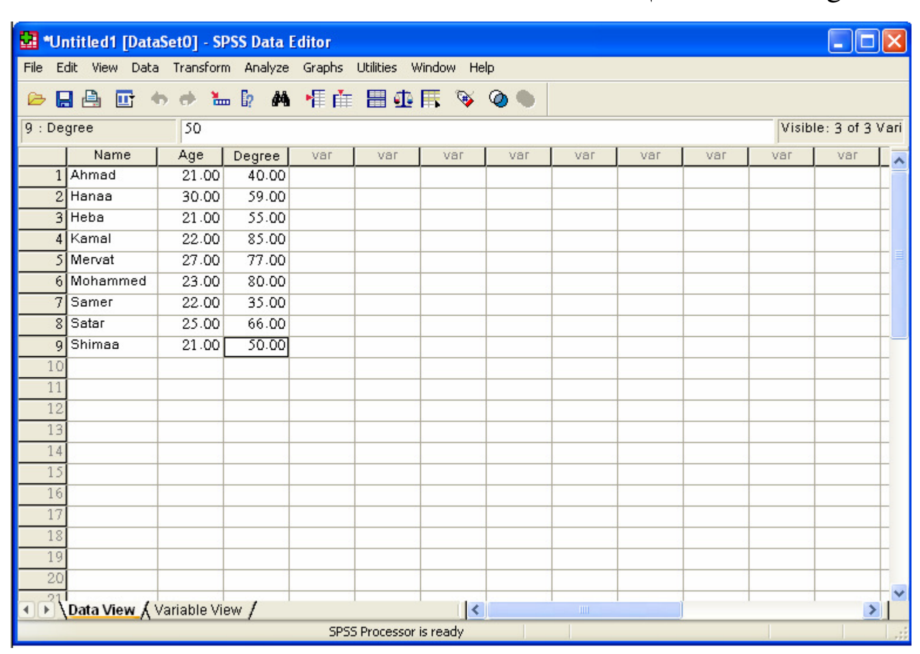
شكل (3-5)

ثم نقوم باختيـار Sort Cases، يفتح الصـندوق الحـواري Sort Cases وبـه أسماء جميع المتغيرات. نختار المتغير الذي سيتم على أساسه الترتيب وهو Name وينقل إلـى المربـع Sort By



الشكل(3-6)

نختار نوع الترتيب من Sort Order هل الترتيب سيكون تصـاعديا Ascending أو تنازليا Descending ثم ننقر على OK فتظهر الشاشة التالية:



شكل (3-7)

نلاحظ بعد تنفيذ الأمر السابق أن البرنامج سيقوم بترتيب البيانات تصـاعديا.

(4-3) دمج الملفات Marge Files :

قد تستدعي الحاجة إجراء عملية ضم ملفات بيانات SPSS و في الحقيقة البرنامج يسمح بإجراء عملية الدمج بين الملفات باستخدام حالات إضافية في نهاية ملفك أو دمج متغيرات إضافية لكل الحالات الموجودة في ملفك:

- الدمج بإضافة حالات جديدة Adding Cases :

هذة العملية تسمح بدمج الملفات التي تحتوى علي نفس المتغيرات ولكن حالات مختلفة:

مثال (1-3) :

نفرض لدينا الملف الأول وبه درجات طالبين لأربع مواد Your data since 1990، وتم حفظ البيانات التاليه في مجلد باسم Our Book:

جدول (2-3)

Name	Math	Chem	Physc	Music
Ahmad	100	90	95	87
Hanaa	95	87	90	85

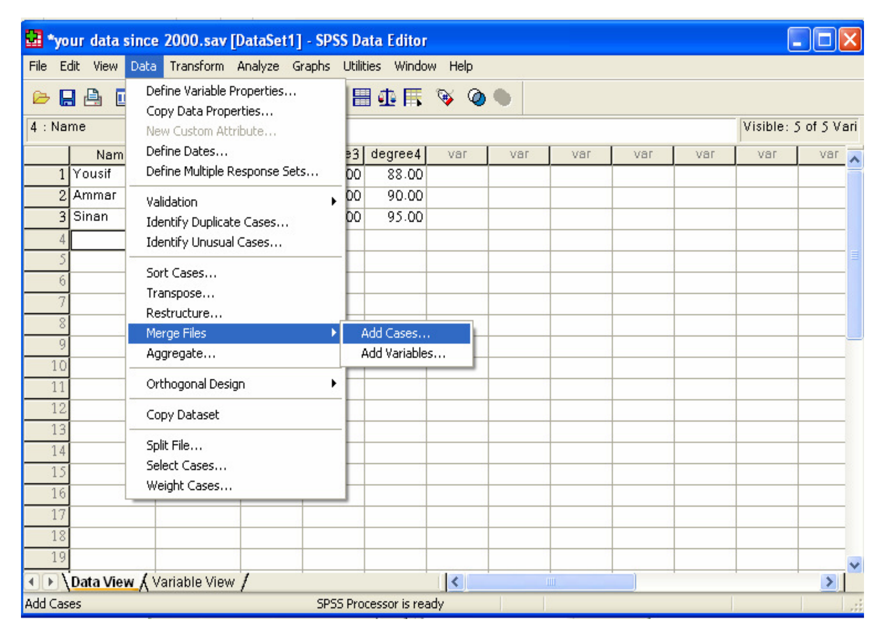
ولدينا الملف الثاني وبه درجات 3 طلاب لأربع مواد Your data since 2000 :

جدول (3-3)

Name	Math	Chem	Physc	Paint
Yousif	85	90	77	88
Ammar	95	83	82	90
Sinan	90	92	86	95

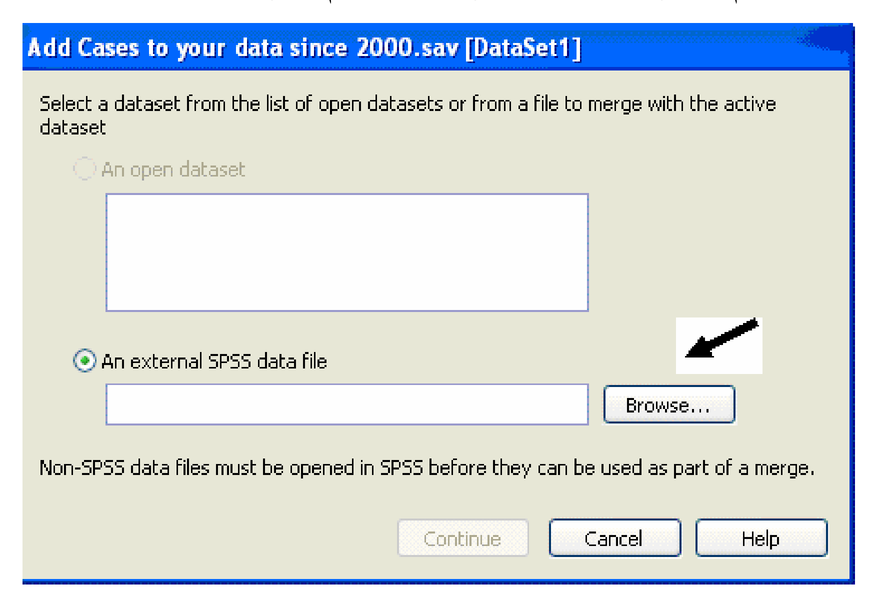
يمكننا من دمج الملفين معاً باستخدام الخطوات التالية:

افتح أحد الملفين في شاشة المحرر وليكن Your data since 2000 ومن قائمة Data اختر الأمر Merge File فيظهر الشكل التالي:



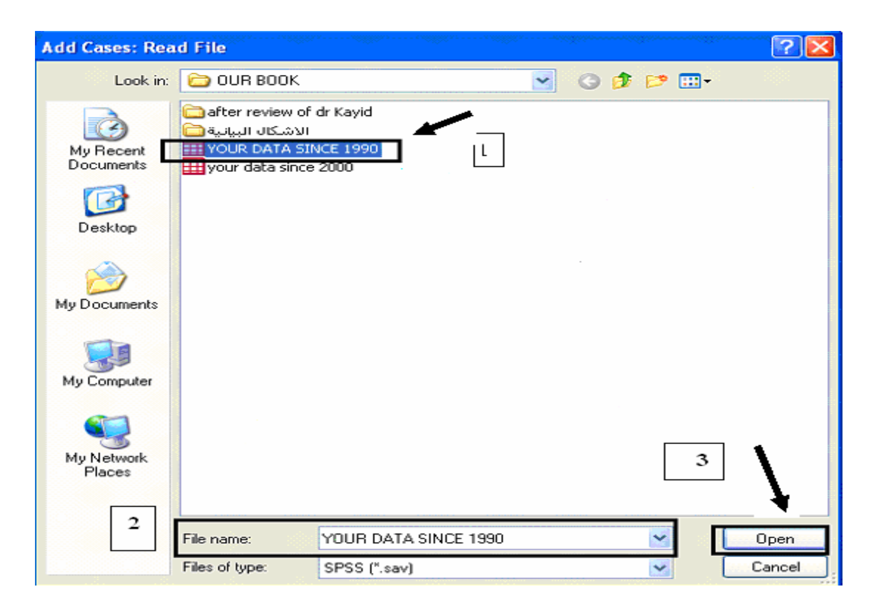
شكل (3-8)

نقوم باختيار add cases بنقر عليها ومن ثم تظهر الشاشة التالية:



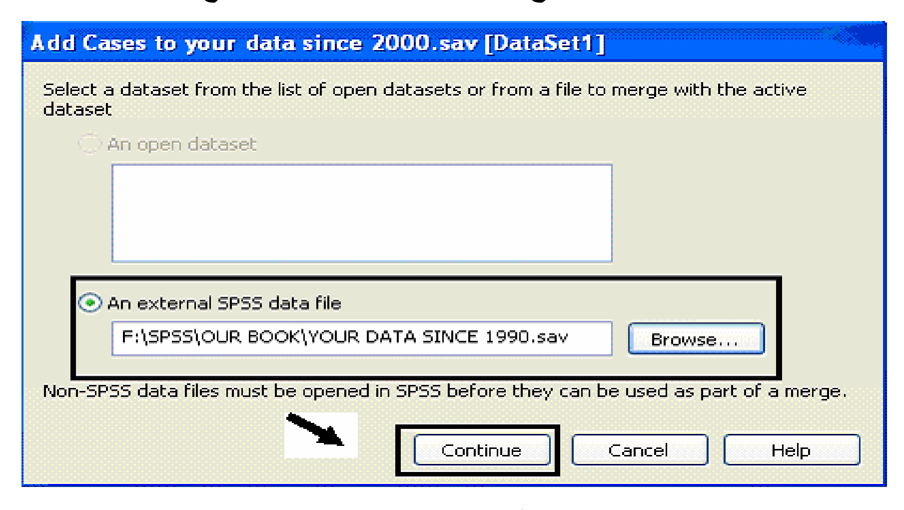
شكل (3-9)

ولأننا ندمج ملف آخر من ملفات SPSS نقوم بتحديد An external SPSS data file وبعدها نقوم بنقر Browse لتحديد مكان الملف الآخر ومن ثم تظهر الشاشة التالية:



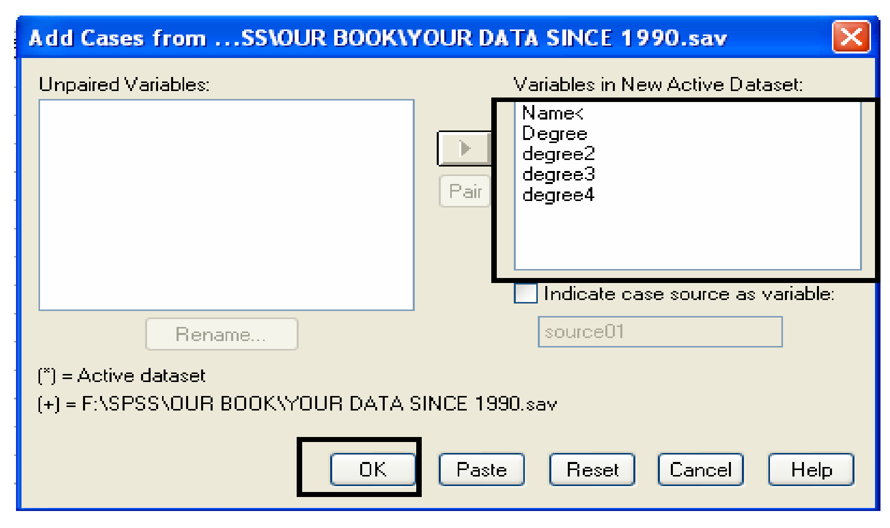
شكل (10-3)

وبعدها تظهر الشاشة التالية موضح بها مكان الملف المراد دمجه مع الملف الآخر:



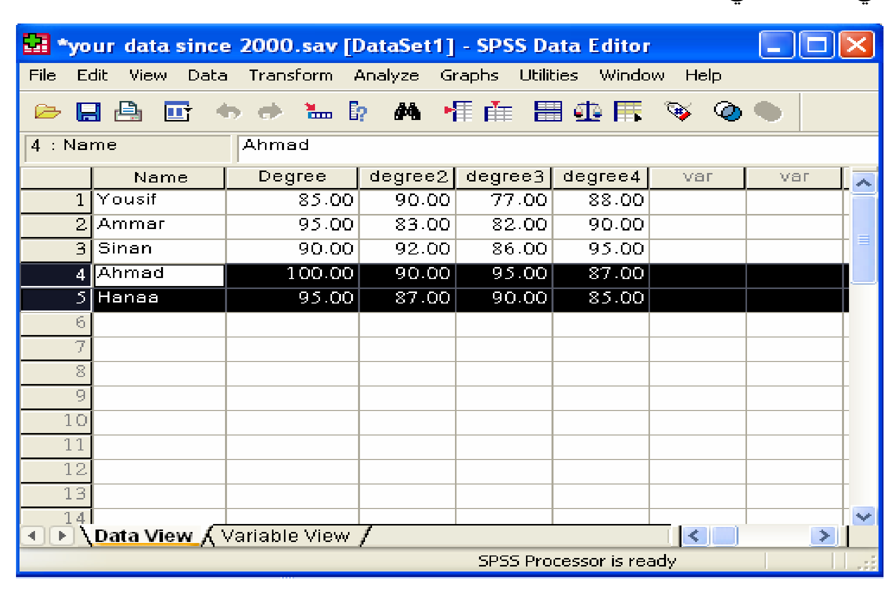
شكل (11-3)

وبنقر continue تظهر الشاشة التالية موضح بها متغيرات ملف your data since 1990 وفي أسفل الشاشة ننقر Ok:



شكل (3-12)

ومن ثم يتم دمج ملف your data since 1990 مع الملف your data since 2000 ونلاحظ أن بيانات الملف المدمج تضاف في نهاية ملف your data since 2000 كما في الشكل الآتي:



شكل (3-13)

- **الدمج بإضافة متغيرات جديدة** Adding Variables :

هذة العملية تسمح بدمج الملفات التي تحتوي على نفس الحالات ولكن متغيرات مختلفة:

مثال (3-2) :

نفرض لدينا الملف الأول وبه بيانات أولية عن عملاء شركة النصر للسيارات:

جدول (3-4)

Name	ID	AGE	Gender
Mervat Mahdy	123	28	1
Ahmad Ali	122	40	2
Mona Mohammed	145	33	1

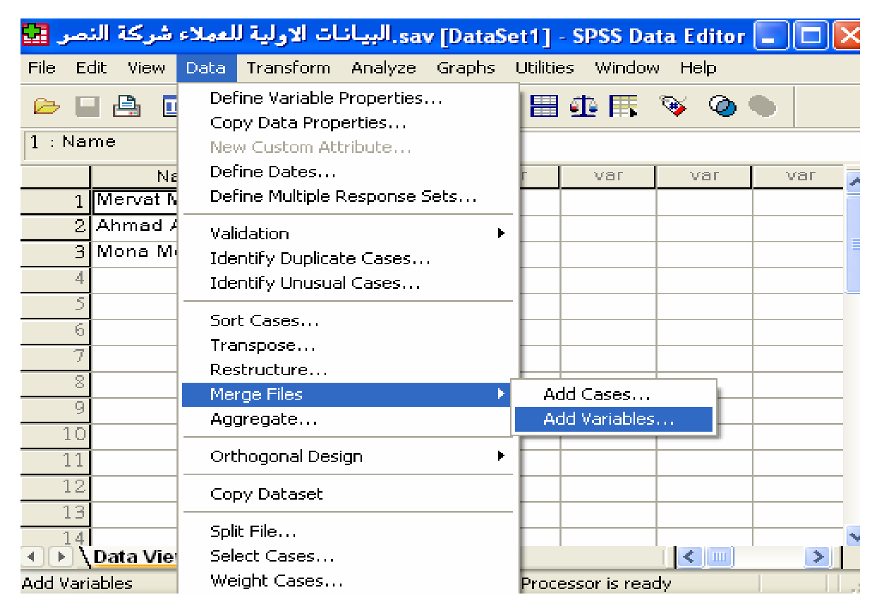
و لدينا الملف الثاني وبه بيانات تفصيلية عن نفس العملاء لنفس الشركة:

جدول (3-5)

Name	Adress	Income
Mervat Mahdy	Cairo	1000
Ahmad Ali	Alex	2000
Mona Mohammed	Giza	1500

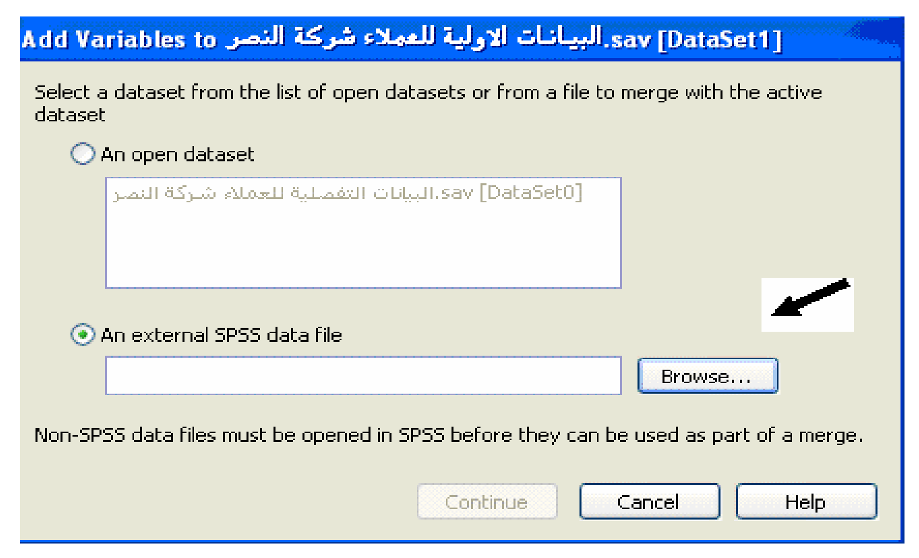
يمكننا من دمج الملفين معاً باستخدام الخطوات التالية:

- ✓ افتح أحد الملفين في شاشة المحرر وليكن البيانات الأولية لعملاء شركة النصر ومن قائمة Data اختر الأمر Merge File فيظهر الشكل التالي:



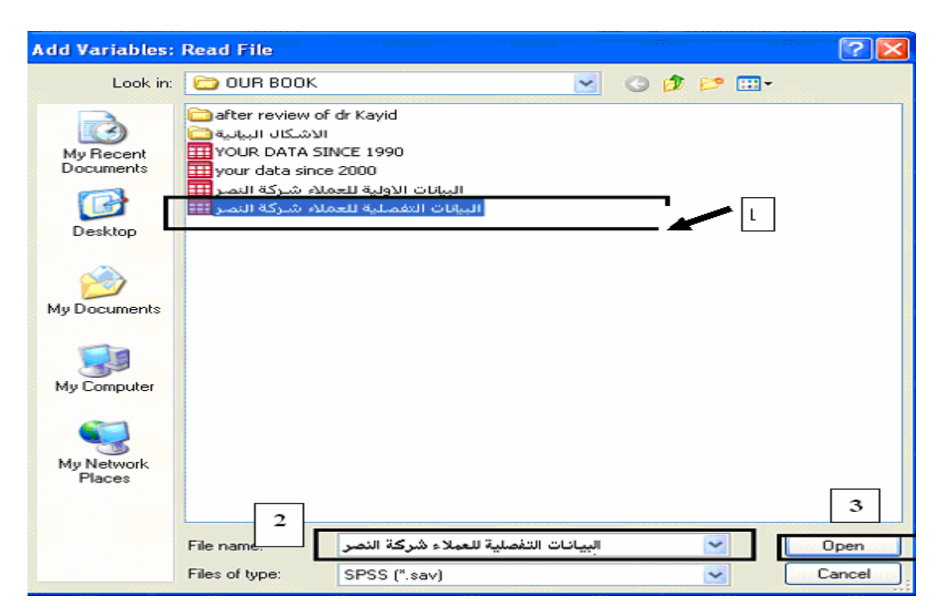
شكل (14-3)

✓ نقوم باختيار add Variables… بنقر عليها ومن ثم تظهر الشاشة التالية:



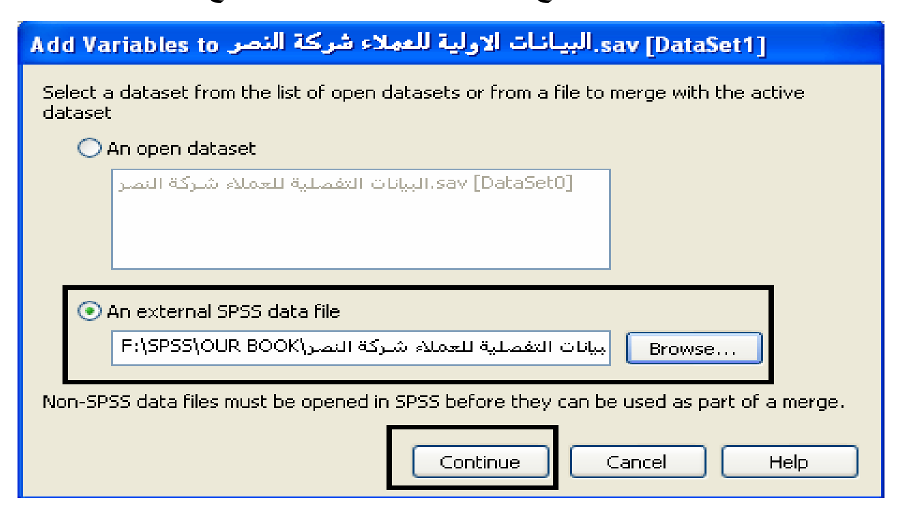
شكل (15-3)

ولأننا ندمج ملف أخر من ملفات SPSS نقوم بتحديد An external SPSS data file وبعدها نقوم بنقر على Browse لتحديد مكان الملف الآخر ومن ثم تظهر الشاشة التالية:



شكل (16-3)

وبعدها تظهر الشاشة التالية موضح بها مكان الملف المراد دمجه مع الملف الآخر:

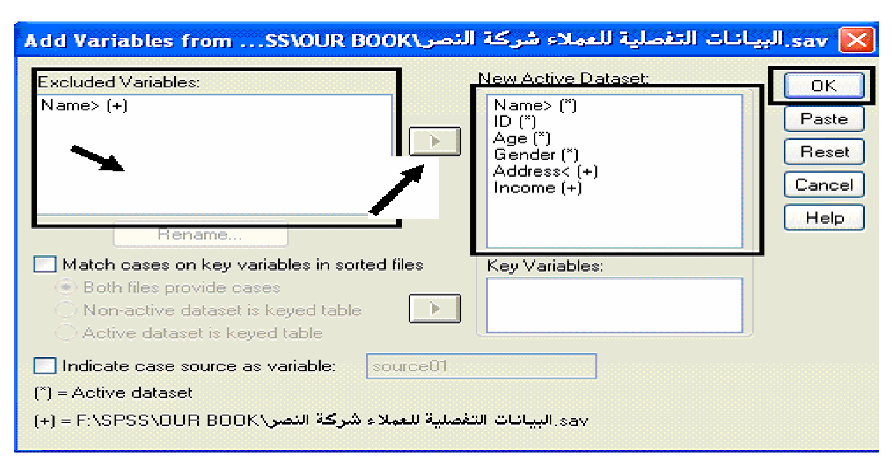


شكل (17-3)

ومع نقر continue تظهر الشاشة التالية وتتضمن ما يلي:

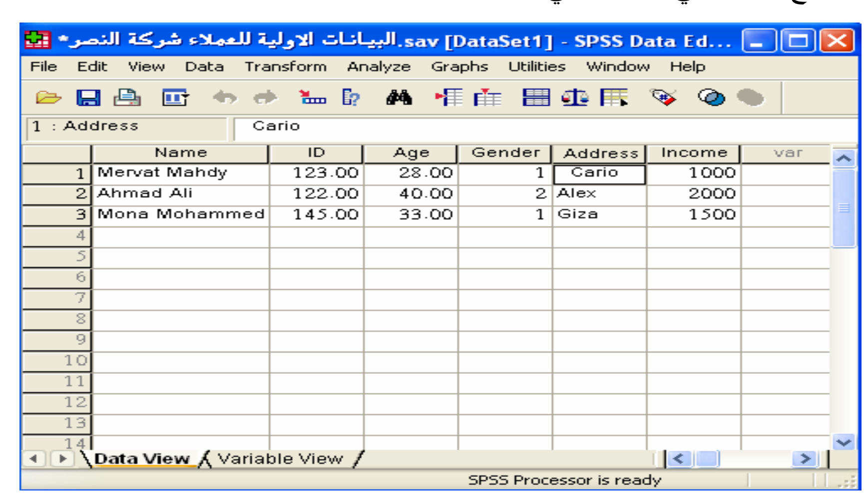
New Active Dataset: وهي قائمة بها جميع المتغيرات في كلا من الملفين التي لا تتشابه من ناحية الاسم في كلا من الملفين.

Excluded Variables: وهي قائمة بأسماء المتغيرات التي تستبعد من الملف المدمج الجديد وعادة ما تتضمن أسماء المتغيرات المشابهة أسمائها وبعد ذلك النقر على Ok:



شكل (3-18)

ومن ثم يتم دمج ملف البيانات الأولية للعملاء شركة النصر مع الملف البيانات التفصيلية للعملاء شركة النصر ونلاحظ أن بيانات الملف البيانات الأولية للعملاء شركة النصر تضاف إليه المتغيرات الجديدة إلى يمين المتغيرات القديمة طالما هو الملف المدمج فيه كما في الشكل الآتي:



شكل (3-19)

(5-3) تحوير ملفات Transpose Files :

نعني بتحوير الملف جعل الصفوف أعمدة والأعمدة صفوف أو بمعنى آخر تحويل الحالات لمتغيرات والمتغيرات لحالات، بشرط أن تكون كلها من نفس النوع.

جدول (6-3)

Name	ID	Age	Gender	Address	Income
Mervat Mahdy	123	28	1	Cairo	1000
Ahmad Ali	122	40	2	Alex	2000
Mona Mohammed	145	33	1	Giza	1500

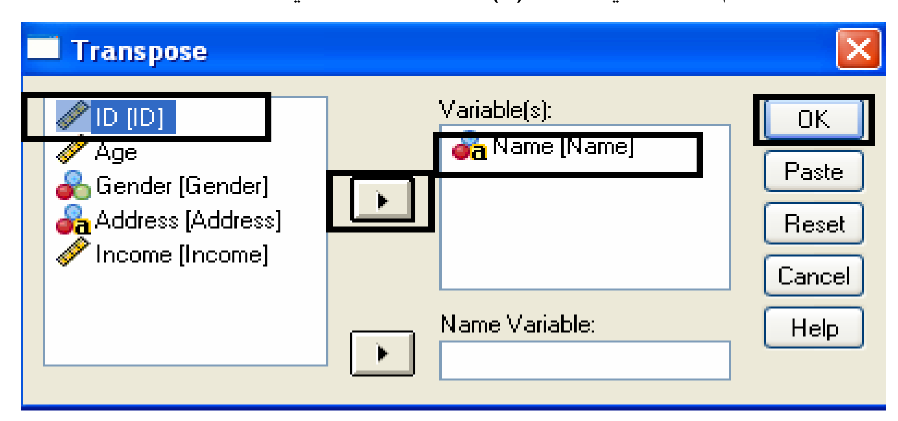
مثال (3-3) :

إذا أردنا تحوير هذا الملف سيكون لدينا ثلاثة متغيرات وحالتين في الملف الجديد.

يمكن إجراء عملية التحوير باستخدام SPSS كالآتي :

- نفتح الملف الذي به البيانات فتكون البيانات في الشكل (19-3) السابق.
- من القائمة Data ننقر على Transpose فيظهر الصندوق الحواري وبه أسماء المتغيرات الموجودة في الملف.
- ننقل المتغيرات إلى مربع Variables بخطوات الآتية:

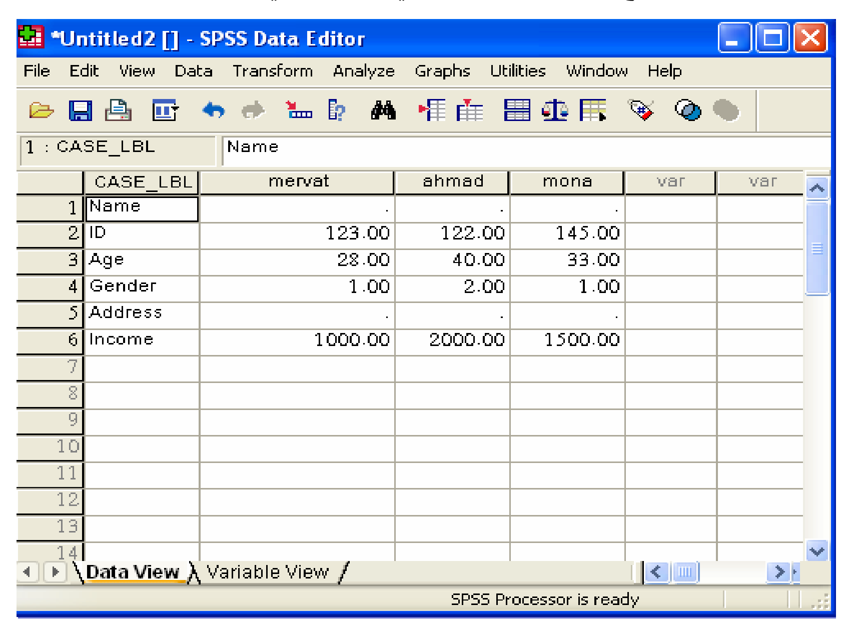
في البداية تحديد المتغير المراد نقله وهو في العمود 1 ثم النقر على مفتاح النقل (2) بعدها يظهر اسم المتغير في العمود (3) وهكذا بالنسبة لباقي المتغيرات.



شكل (20-3)

✓ ننقر Ok لنحصل على الملف المحور.

فيصبح شكل ملف البيانات في الشكل الآتي:



	CASE_LBL	mervat	ahmad	mona	var	var
1	Name	.	.	.		
2	ID	123.00	122.00	145.00		
3	Age	28.00	40.00	33.00		
4	Gender	1.00	2.00	1.00		
5	Address	.	.	.		
6	Income	1000.00	2000.00	1500.00		
7						
8						
9						
10						
11						
12						
13						
14						

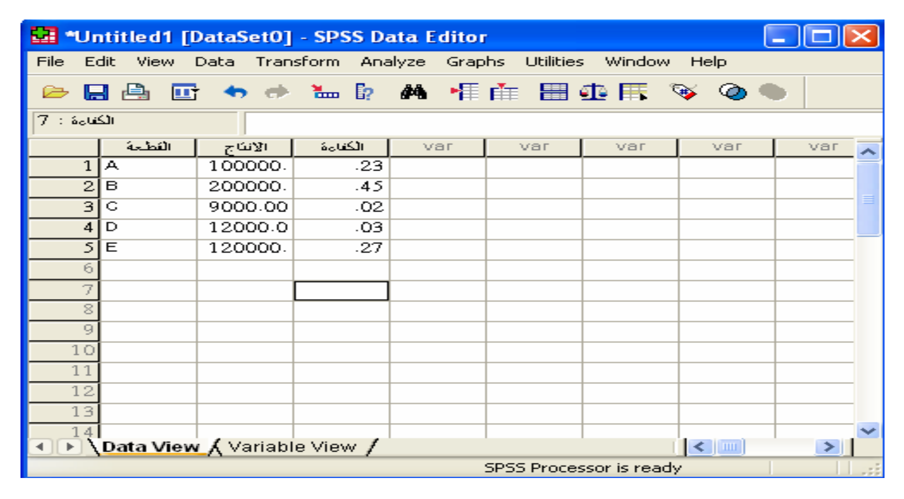
شكل (3-21)

(3-6) **تلخيص الحالات** Aggregate :

يستعمل هذا الأمر لتلخيص المعلومات المتعلقة بمجموعة من الحالات cases في حالة تجميعية واحدة وتكوين ملف تجميعي جديد ويمكن أن نجري تلخيص للحالات باستخدام البرنامج كما في المثال التالي:

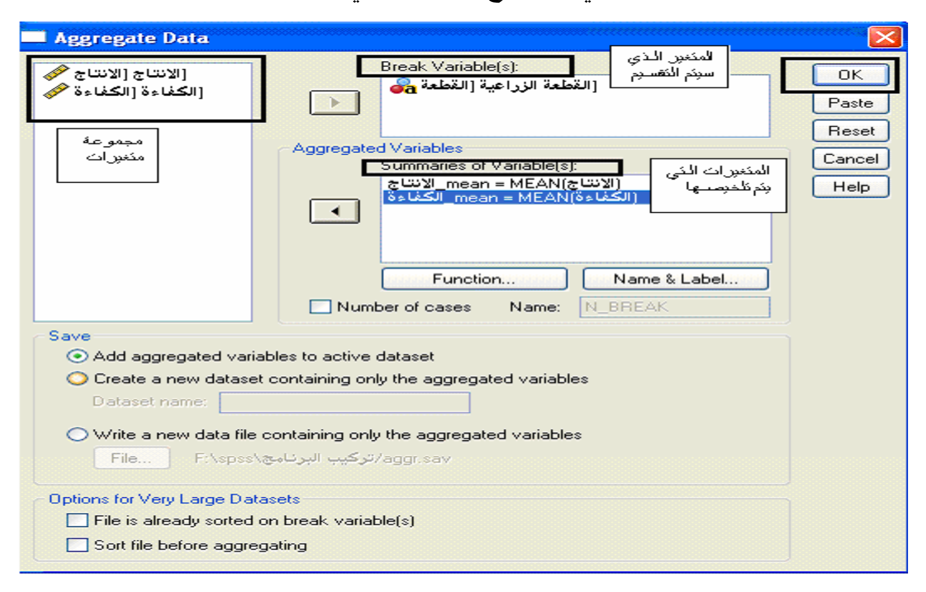
مثال (3-4):

الملف الآتي يوضح بيان بإنتاجية 5 قطع زراعية معينة تنتج محصول القمح ونسبة كفاءة كل قطعة زراعية كما يلي:



شكل (3-22)

من القائمة Data انقر الأمر Aggregate يظهر الصندوق Aggregate Data التالي الموضح بالشكل التالي:



الشكل (3-23)

⌖ من قائمة المتغيرات على اليسار اختر المتغير الذي سيتم على أساسه التقسيم وانقله إلى Break Variable ويتم النقل عن طريق زر الانتقال الموجود بجوار Break Variable من ناحية اليسار الظاهر بشكل ، ونلاحظ أنه عندما ننقل هذا المتغير فإنه يحذف من عمود مجموعة المتغيرات.

- قم بنقل المتغير (أو المتغيرات) المطلوب اجراء العمليات الإحصائية لها بناء على متغير التقسيم إلى Summaries of variable(s). لاحظ أننا يمكننا الرجوع عن الاختيار باستخدام الزر الآتي الموجود بجوار Summaries of variable(s) من ناحية اليسار الظاهر بشكل .
- انقر على الأمر Function فيظهر لك الصندوق Aggregate Function: Aggregate Data وهو يحتوي على عمليات إحصائية كثيرة ولاحظ انه لا يحق للمستخدم أن يختار أكثر من عملية إحصائية في نفس الوقت (انظر الشكل 3-24).

الشكل (3-24)

- انقر على الأمر Continue للعودة للصندوق الأصلي .
- انقر على الأمر Ok للتنفيذ.

لاحظ انه تم إيجاد الوسط الحسابي كما طلبنا في الأمر Aggregate Function.

(3-7) فصل الملفات Split Files :

يستعمل هذا الأمر لغرض تجزئة (فصل) ملف البيانات لأغراض التحليل الإحصائي .

مثال (3-5) :

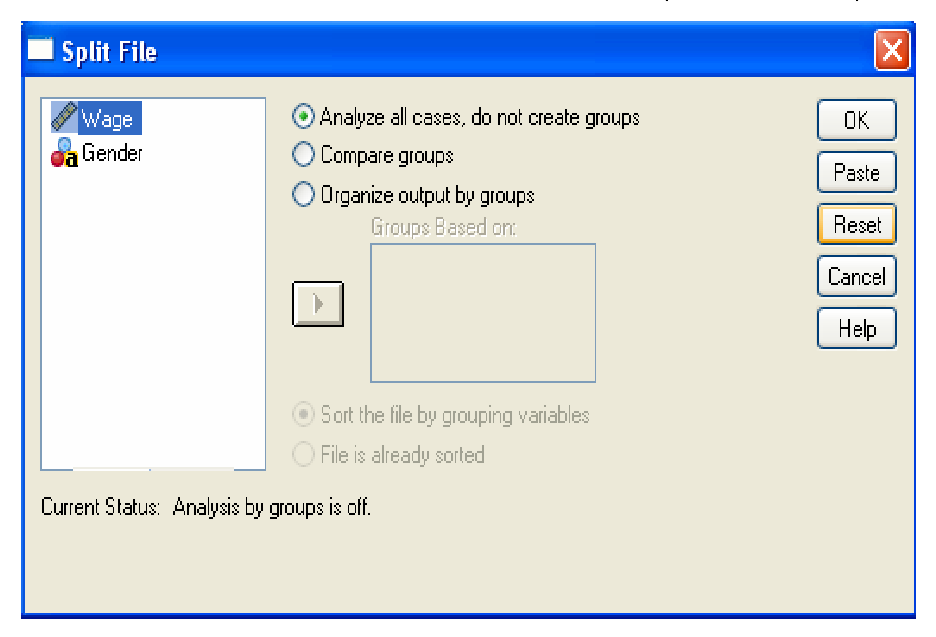
الجدول التالي يمثل رواتب مجموعة من الأشخاص حسب الجنس.

جدول (3-7)

Wage	Gender
60	M
30	F
70	M
35	F
65	M
40	F

لتجزئة الملف إلى جزأين الأول يمثل رواتب الذكور M والثاني يمثل رواتب الإناث F نتبع الخطوات التالية :

⌖ من شريط القوائم اختر Data ثم Split File سوف يظهر صندوق Split File (انظر الشكل 3-25) .



الشكل (3-25)

الصندوق يحتوي على جميع المتغيرات الموجودة في الملف وهناك عدة اختيارات وهي:

- Do Not Create Groups،Analyze Cases عدم تجزئة الملف .
- Compare Groups تجزئة الملف حسب فئات متغير معين .
- Organize output by Groups ترتيب بيانات الملف بالنسبة لمتغير معين .

يوجد بعد ذلك خيارين مطلوب اختيار أحدهما (هذه الخيارات غير نشطة في حالة التعامل مع بيانات الملف دون تقسيم) وهما:

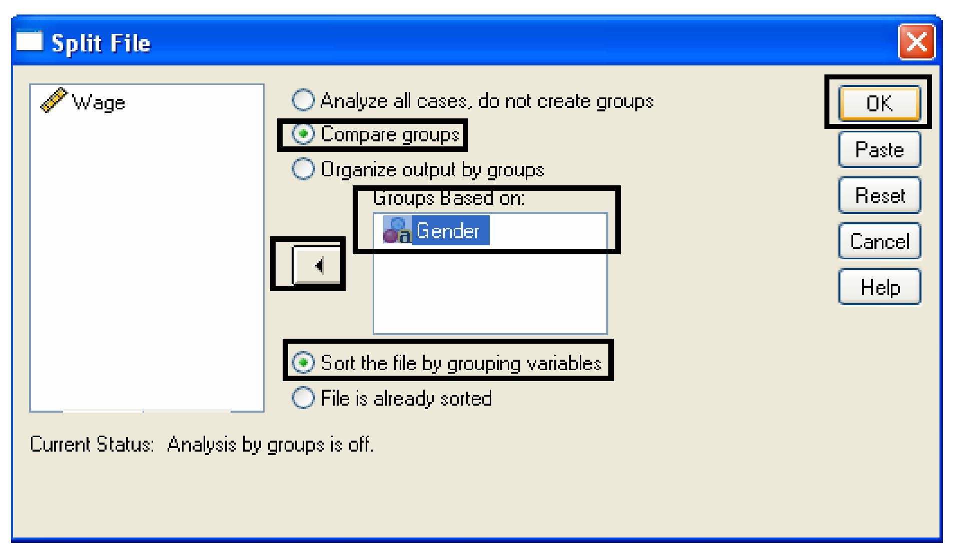
أ - ترتيب الملف طبقا للمتغير الذي تم على أساسه الفصل .

ب - لا داعي لترتيب المتغيرات في الملف .

ملاحظة : الخياران أ ، ب تكون نشطة عند اختيار الخيارين 2 ، 3 .

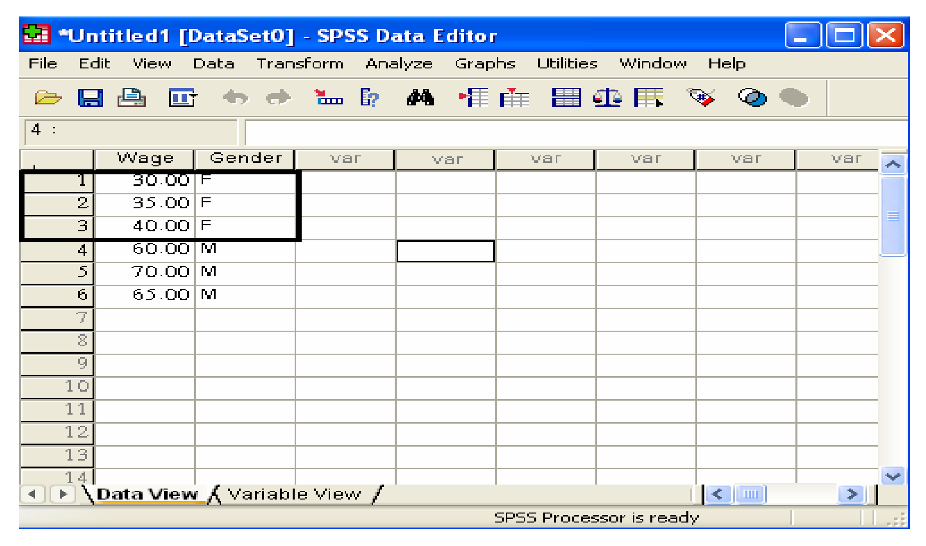
انقر على الخيار الثاني وانقل المتغير الذي على أساسه سيتم التقسيم وهو Gender إلى الخانة المخصصة لذلك وهي Grouped Based On .

- انقر على أي من الخيارين وهما الأول لترتيب الملف طبقا لمتغير التقسيم Sort The File By Grouping Variables والثاني أن الملف غير مرتب Sort the file by grouping variables (انظر شكل 3-26).



الشكل (3-26)

- انقر على Ok يتم تنفيذ الأمر فيظهر ملف النتيجة كما يلي:



الشكل (3-27)

(3-8) ترجيح حالات Weight Cases:

يتيح هذا الأمر إمكانية إعطاء أوزان لحالات Cases ملف معين نظرا لاختلافها من ناحية الأهمية و هي نفس الوظيفة التي تقوم بها الأيقونة ⚖ في شريط الأدوات .

مثال (3-6):

القيم التالية تمثل نتائج أحد الطلاب في اختبار مادة الإحصاء علما أن لكل اختبار وزنا (أهمية) مختلفة:

جدول (3-8)

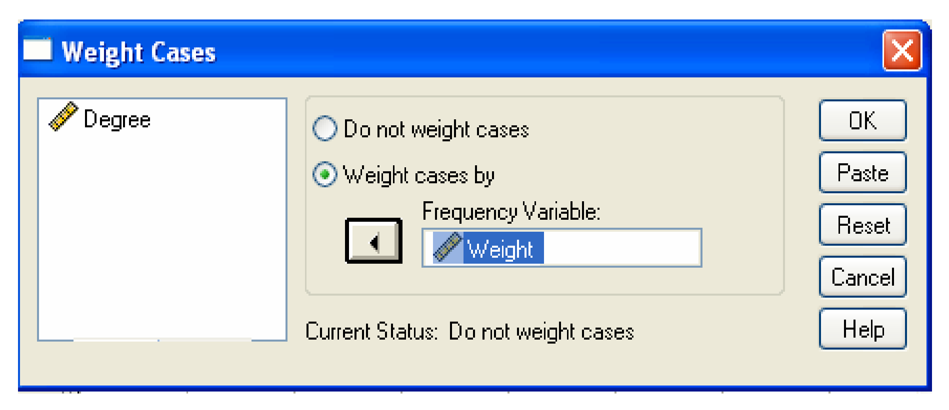
Test	Degree	Weight
A	100	50
B	60	35
C	75	5
D	55	15

المطلوب حساب الوسط الحسابي المرجح للامتحانات الأربعة؟

الحل:

لحساب الوسط الحسابي المرجح بالمتغير Weight للامتحانات الأربعة نتبع الخطوات التالية:

⌖ من شريط القوائم Data نختار Weight Cases الذي نقوم بترتيبه (الشكل 3-28):



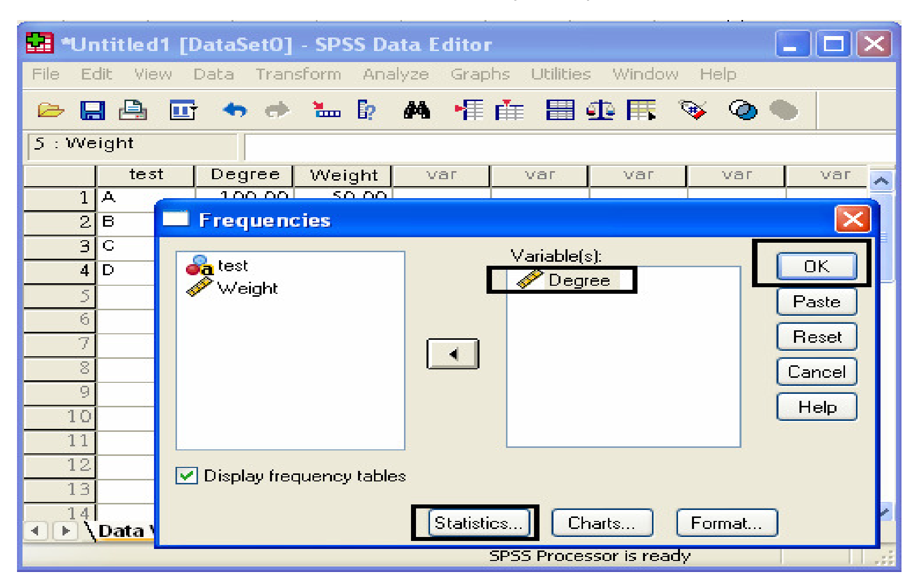
الشكل(3-28)

⌖ عند نقر زر Ok يتم وزن حالات الملف بالمتغير Weight (علما أنـه لا يلاحظ أي تغير في شاشة Data Editor).

لحساب الوسط الحسابي نختار من شريط القوائم :

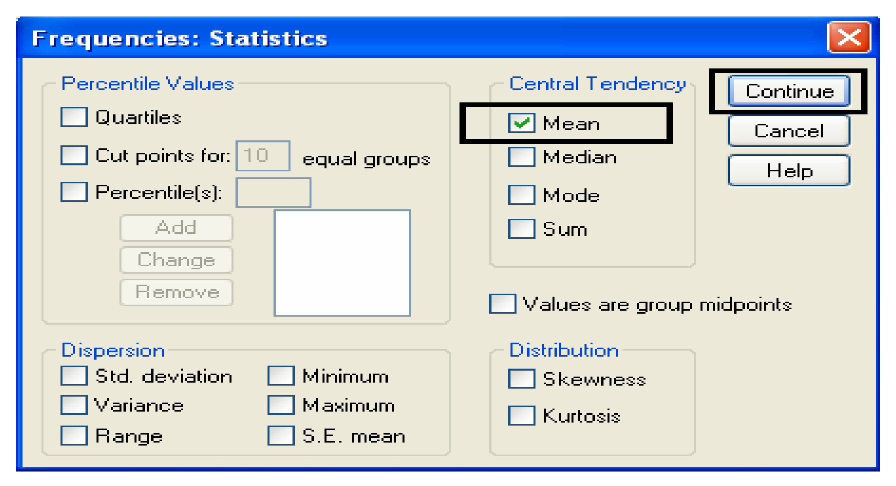
Analyze ➡ Descriptive Statistics ➡ Frequencies

انظر الشكل (3-29):



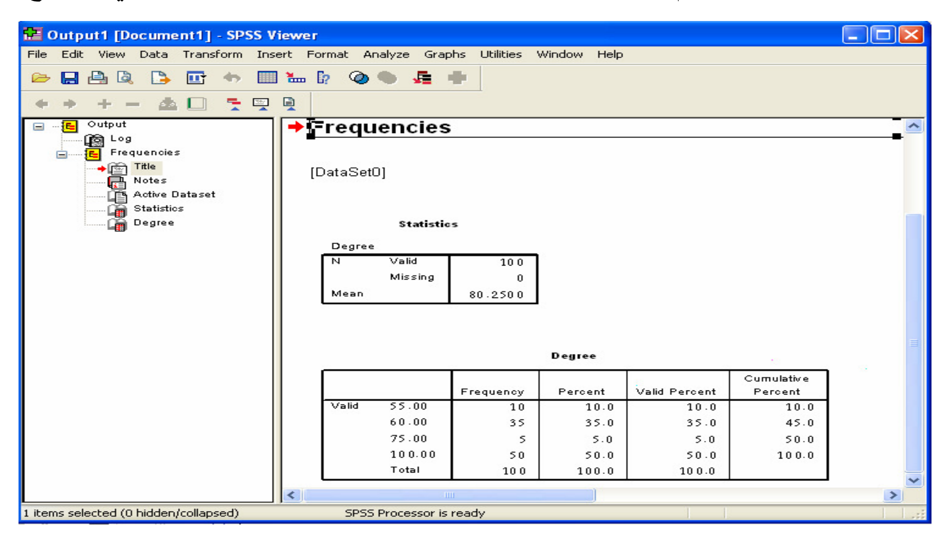
الشكل (3-29)

⌖ نقوم بتحديد المتغير المراد حساب الوسط الحسابي المرجح له ويتم نقله عن طريق زر النقل ويمكننا الرجوع عن الاختيار باستخدام زر الرجوع وباختيار Statistics نقوم بتأشير الخيار Mean كما في الشكل الآتي:



الشكل (3-30)

⌖ وبنقر continue ثم Ok تظهر شاشة المخرجات الآتية بها الوسط الحسابي المرجح:



الشكل (3-31)

ملاحظة هامة:

✓ لإلغاء ترجيح حالات الملف نقوم بتأشير الخيار Do Not Weight Cases في صندوق حوار Weight Cases.

الفصل الرابع

إنشاء متغيرات جديدة

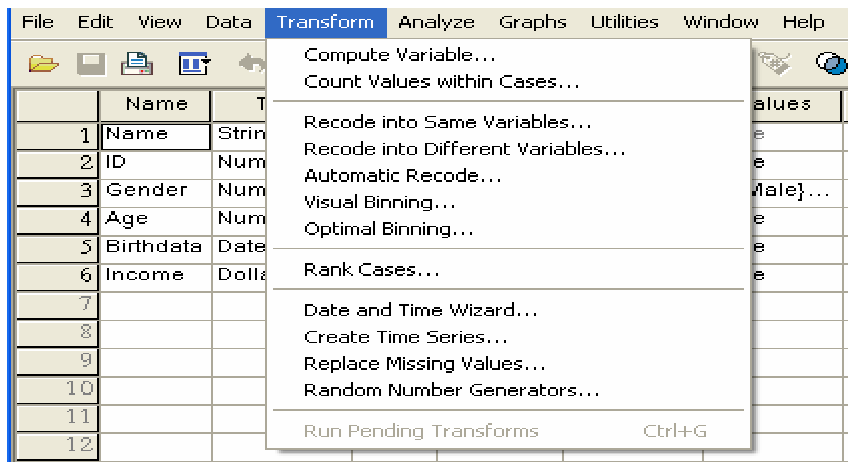
Create New Variables

(1-4) مقدمة:

يمكن إجراء تحويلات على البيانات الأصلية ما بين تجميع بيانات في فئات معينة إلى تكوين متغيرات جديدة بالاعتماد على معادلات وصيغ شرطية فمثلا إذا كان لدينا بيانات استهلاك ودخول أفراد مجتمع معين فيمكننا من حساب حجم الادخار في هذا المجتمع عن طريق إنشاء متغير جديد عبارة عن الفرق بين الدخل والاستهلاك ليمثل الادخار تمهيداً لاستعماله في حساب حجم الاستثمارات في المجتمع، والبيانات التي لا تتبع التوزيع الطبيعي يمكن تحويلها عن طريق هذا الفصل إلى بيانات طبيعية. ونجد أنه في بعض الأحيان يصعب التعامل مع البيانات اللفظية ويكون التعامل أسهل عند تكويد البيانات اللفظية وتحويلها إلى أكواد رقمية ويتم ذلك عن طريق التحويلات. هذا الباب سيناقش كيفية التعامل مع القائمة Transformations.

(2-4) قائمة Transform:

تضم قائمة Transform الأوامر التالية:



شكل (1-4)

(2-4) استخدام العمليات الرياضية لإنشاء متغيرات مستمرة جديدة/الأمر Compute:

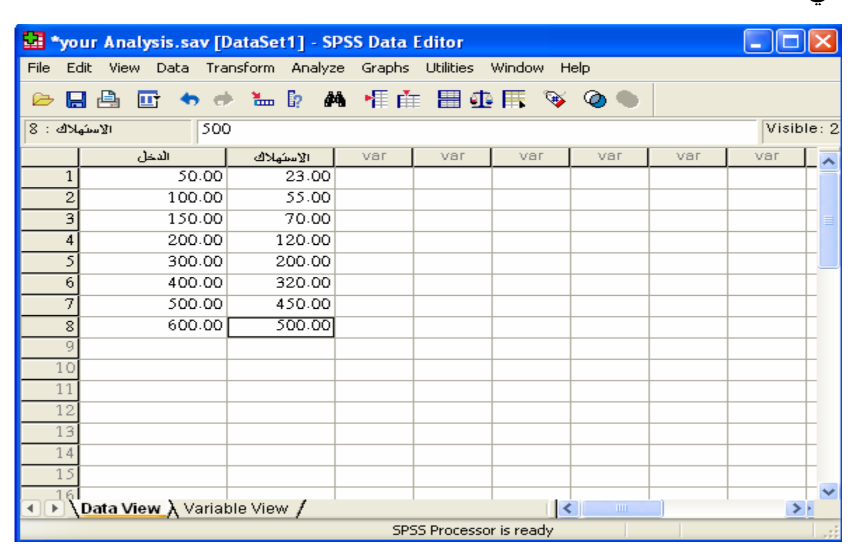
يتيح هذا الأمر إمكانية حساب متغيرات جديدة باستخدام أكثر من دالة تتضمن (دوال إحصائية ، توزيعات احتمالية)، أو تحويل متغيرات موجودة إلى متغيرات أخرى.

⌖ **العمليات البسيطة:**

يمكننا من استعمال بعض الدوال البسيطة مثل الجمع والطرح والضرب والقسمة والأس وهكذا في إجراء تحويل المتغيرات كما في المثال التالي.

مثال (1-4):

نفترض أننا نريد حساب حجم الادخار للبيانات التالية الموجودة في Data view التالي:

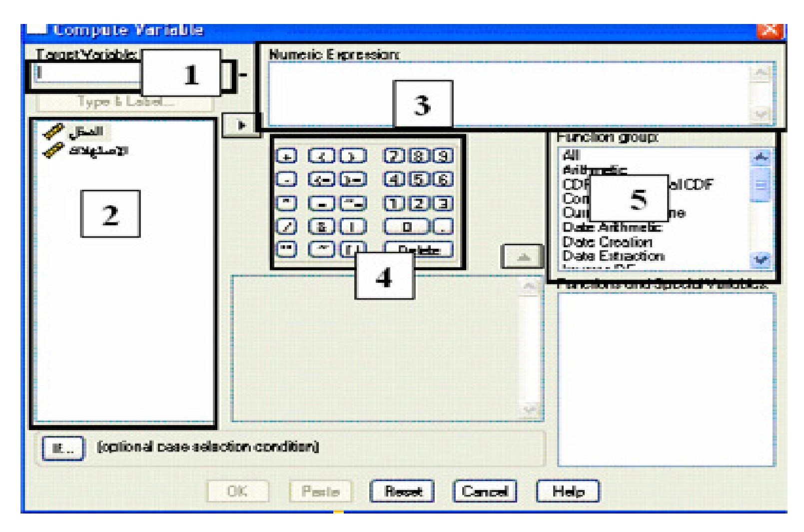


شكل (2-4)

فنجد أن: الادخار=الدخل - الاستهلاك

للحصول على عمود الادخار فإننا نقوم بطرح عمود الاستهلاك من عمود الدخل باستعمال الخطوات التالية:

⌖ من قائمة Transform نختار Compute Variable فيظهر الشكل الآتي:



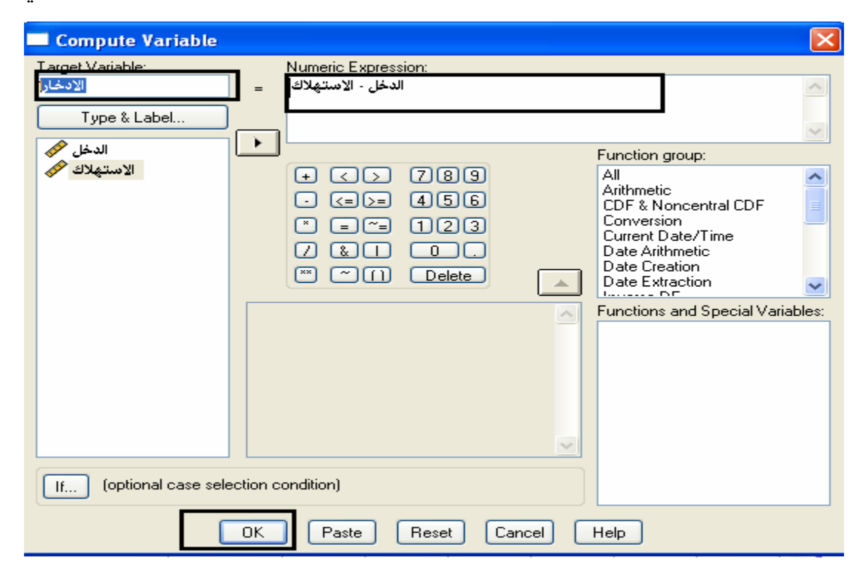
شكل (4-3)

- في المنطقة رقم "1" نضع اسم المتغير الجديد.
- في المنطقة رقم "2" وهي المتغيرات المتاح الاختيار منها لإجراء العمليات الحسابية عليها وإنشاء متغير منها ونلاحظ أننا لا نمتلك سوا متغيرين اثنين وهو الاستهلاك والدخل وظاهرين في الشكل (4-3).
- في المنطقة رقم "3" وهي المنطقة الذي يظهر بها شكل العمليات الحسابية المطلوبة.
- في المنطقة رقم "4" هي منطقة تتيح لوحة مفاتيح للأرقام وبعض العمليات الحسابية البسيطة. والجدول التالي يوضح ملخص لبعض العمليات الحسابية البسيطة المتوفرة في هذة المنطقة:

جدول (4-1)

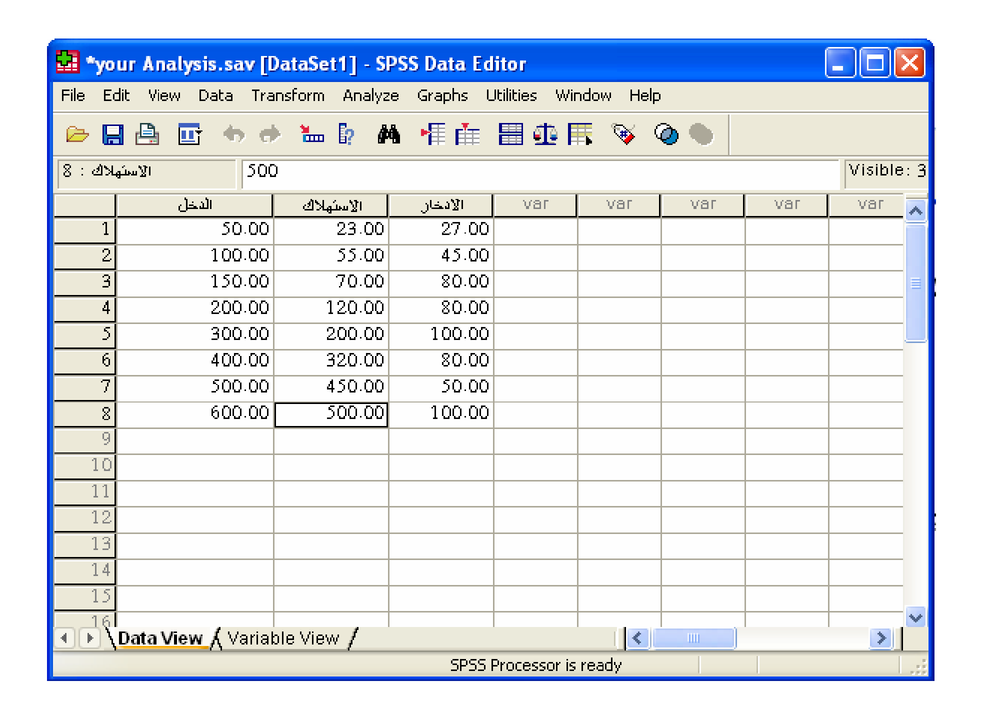
العملية الحسابية	الرمز
الجمع	+
الطرح	-
القسمة	/
الضرب	*
الأس	"**" أو "^"

- في المنطقة رقم "5" هي منطقة الدوال الأكثر تعقيداً من الدوال الموجودة في المنطقة رقم "4".
- ففي هذا المثال نقوم بكتابة كلمة "ادخار" في المنطقة رقم "1" ليكون اسم المتغير الجديد.
- وبعدها نختار المتغير الدخل من المنطقة رقم "2" عن طريق تحديد المتغير والنقر علي زر الانتقال ▶ ونلاحظ أن متغير الدخل ظهر بعد ذلك في المنطقة رقم "3" تمهيداً لإجراء العمليات الحسابية عليه.
- وبعدها نختار علامة الطرح من المنطقة رقم 4 وهي بالشكل الآتي [-] ونلاحظ أن بعد الاختيار فإن تلك العلامة تظهر في المنطقة رقم "3" بعد متغير الدخل.
- وبعدها نختار مرة أخرى متغير الاستهلاك من المنطقة "2". فيظهر الشكل الآتي:



شكل (4-4)

- ننقر على Ok فتظهر شاشة data view كالتالي:



شكل (4-5)

- **استخدام دوال SPSS لإنشاء متغيرات جديدة:**

يوفر لنا البرنامج مجموعة من الدوال ليست فقط الرياضية ولكن دوال إحصائية ومنطقية أيضا. ففي المثال الآتي يوضح كيف يمكننا استخدام الدوال الإحصائية في إنشاء متغيرات جديدة:

مثال (4-2):

الجدول التالي يمثل المتغيرين X_1، X_2 اللذين تم إدخالها إلى Data Editor كما يلي:

جدول (4-2)

X_1	X_2
60	90
87	88
70	43
90	80
57	55

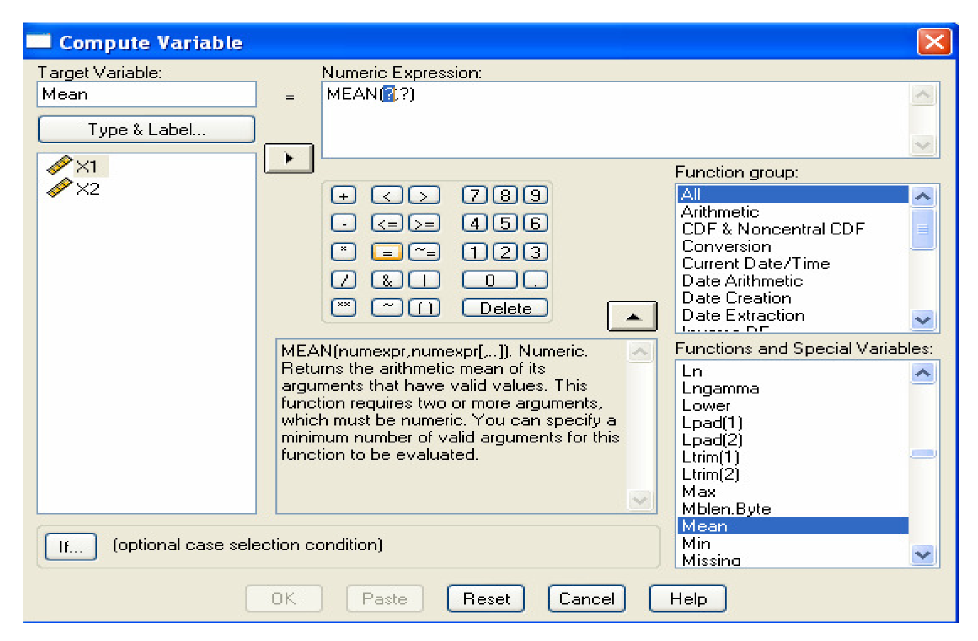
73	47
95	90
66	50
40	55
55	80
85	75
88	86
35	70

المطلــوب حســاب الوســط الحســابي Mean للمتغيــرين X_1 و X_2 فــي حالــة كــون X_1،$X_2 \geq 50$.

الحل:

لتنفيذ ذلك نتبع الخطوات التالية:

- مـن شـريط القـوائم نختـار Transform ثـم ننقـر Compute فيظهـر صـندوق حـوار Comput Variable انظـر الشـكل (6-4) ونكتـب اسـم المتغيـر الجديـد وهو الوسـط الحسـابي (mean) لكـلاً من X_1 and X_2 يلـي ذلك اختيـار الدالـة Mean مـن منطقـة الـدوال المتاحـة فتظهـر فـي منطقـة العمليـات الحسـابية وبجوارها علامتين استفهام العلامـة الأولـي متاحـة لكتابـة المتغيـر الأول والثـاني متاحة للكتابة بدلاً منها المتغير الثاني ويمكن كتابة أكثر مـن متغير مفصـول بـين كل متغير والآخر بفصلة:



شكل (4-6)

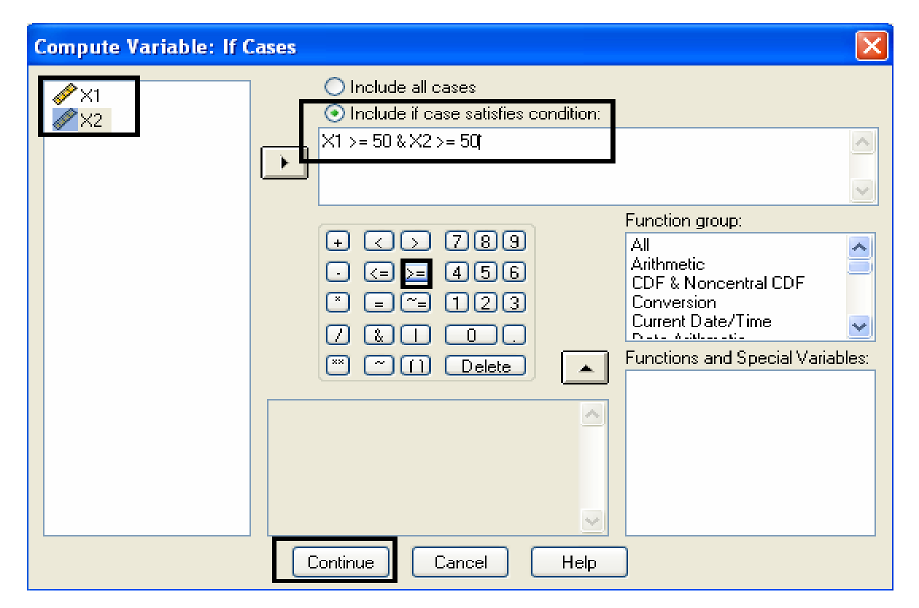
فنقوم بكتابة X_1 بدلاً من علامة الاستفهام الأولى "؟" وكتابة X_2 بدلاً من علامة الاستفهام الثانية "؟"

وقد أجرينا العمليات التالية:

- تحديد اسم المتغير Target Variable الذي هو عبارة عن الوسط الحسابي لـ X_1،X_2 وهـو X_3.
- اختيار الدالة وهي Mean من قائمة Function.
- واضح في صندوق الحوار أعلاه.
- بعد الانتهاء انقر IF في صندوق حوار شكل (4-6) حيث لدينا شرط في المتغير الجديد وهو أن يكون $50 \geq X_1$،X_2 فيظهر صندوق حوار If Cases كما بالشكل (4-7):

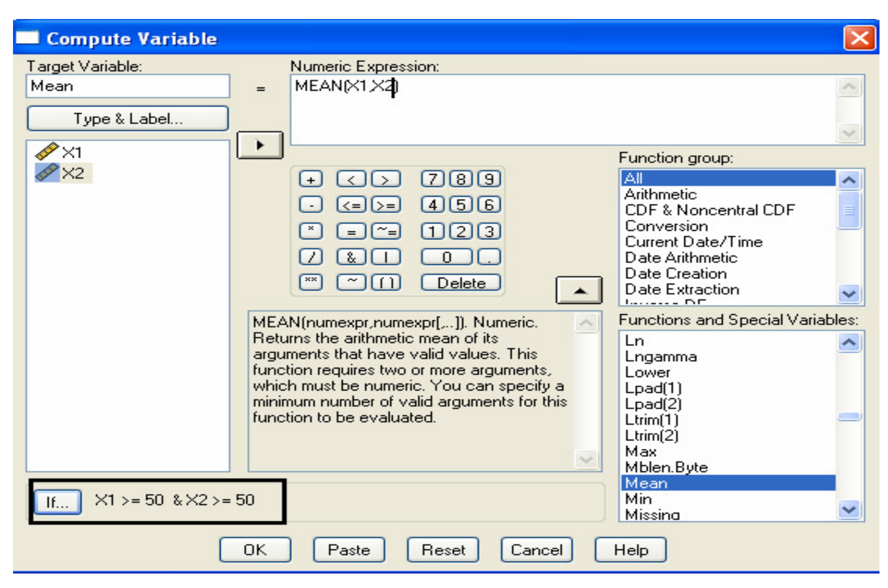
✓ لاختيار كافة الحالات انقر Include All Cases.

✓ لاختيار جزء من الحالات انقر Include If Cases Satisfies Condition. بما أننا نريد اختيار حالات نقوم بنقر الجزء الأخير. ونقوم بكتابة الشرط في شاشة الشرط كما في الشكل الآتي:



الشكل (7-4)

⌖ انقر Continue في صندوق حوار if cases للعودة إلى صندوق الحوار الأصلي ونلاحظ أن الشرط قد تم كتابته بجوار أيقونة If كما في الشكل الآتي:



الشكل (8-4)

⌖ انفر زر ok في صندوق حوار Compute Variable.

يتم الحصول على نتائج حساب المتوسط حيث يتم إضافة متغير آخر هو X_3 إلى Data Editor كما يلي:

X1	X2	Mean
60.00	90.00	75.00
87.00	88.00	87.50
70.00	43.00	.
90.00	80.00	85.00
57.00	55.00	56.00
73.00	47.00	.
95.00	90.00	92.50
66.00	50.00	58.00
40.00	55.00	.
55.00	80.00	67.50
85.00	75.00	80.00
88.00	86.00	87.00
35.00	70.00	.

الشكل (4-9)

الأمر Recode:

في كثير من الأحيان قد نحتاج إلى عملية التكويد (أو الترميز) وذلك بتعديل قيم البيانات وتغييرها من بيانات لفظية إلى بيانات رقمية بغرض تسهيل التعامل معها، ويستفاد منه أيضا عمل الفئات. يوجد عدة حالات يمكن فيها استخدام عملية التكويد منها الآتي:

⌖ دمج بيانات كثيرة لمتغير مستمر في جدول تكراري ليسهل التعامل معها. للتعرف على تكويد البيانات المستمرة نفرض أن لدينا المثال التالي :

مثال (4-3):

لدينا المتغير Salary (X) (الأجر) يأخذ القيم التالية:

100, 150, 200, 250, 300, 500, 270, 300, 350, 700, 350, 400, 550, 600, 850

وقد تم ادخال قيم المتغير في شاشة Data Editor. وتم تخصيص كود لكل قيمة من قيم المتغير حسب الترتيب التالي:

جدول (3-4)

الفئة	الكود
150 فأقل	1
300-151	2
450-301	3
600-451	4
601 فأكثر	5

نرغب في ترميز المتغير Salary حسب الفئات المذكورة وخزن الرموز في متغير مختلف.

الحل:

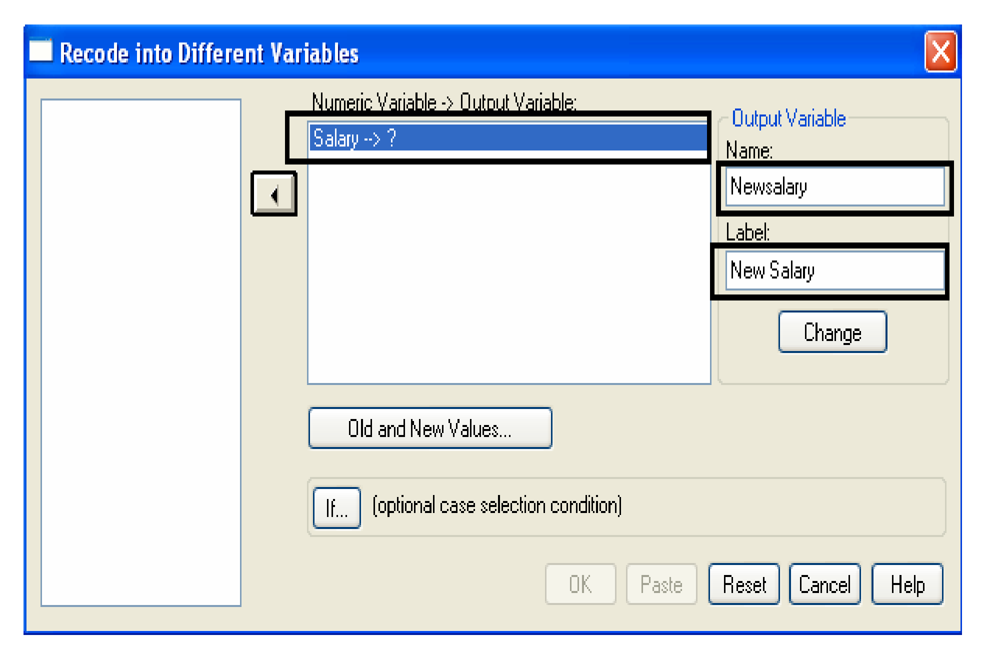
لتنفيذ ذلك نتبع الخطوات التالية :

- من القائمة Transforms عند النقر عليها سوف تظهر شاشة (4-1) وهنا يوجد في هذا المجال اختيارين:

* Recode into the Same Variable وهذا يعني أن نجري العملية ونضع القيم الجديدة فوق القيم القديمة للمتغير .

*Recode into Different Variable بينما يعني هذا أن نضع القيم الجديدة تحت اسم متغير جديد.

- ✓ عند الاختيار الأول تتم العملية تلقائياً، أما عند الثاني فسيطلب البرنامج معلومات عن المتغير الجديد.
- ✓ عند اختيار الأمر الثاني سيفتح الصندوق الحواري Recode Into Different Variable ننقل المتغير المراد تكويده إلى Input Variable وليكن Salary (الأجر)، ثم نحدد اسم المتغير الجديد في المكان المسمى Name وليكن Y انظر الصندوق الحواري الشكل (4-10):



الشكل (4-10)

⌖ تبدأ عملية التكويد بالنقر على Old and new values فيفتح الصندوق الحواري

Recode Into Different Variable: Old And New Values انظر الشكل (4-11):

✓ فإذا كنا نريد تكويد قيمة



فإننا نقوم بنقر تحت الأمر Old Values ولكننا هنا نريد تكويد فئات لها أكثر من نوع.

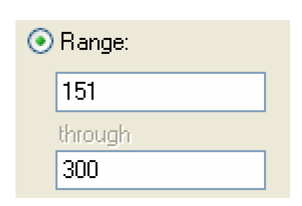
✓ أول فئة لدينا هي 150فأقل فإننا نختار



تحت الأمر Old Values ثم يلي ذلك أن ندخل الترميز "1" في المستطيل الموجود تحت كلمة New value ويكون بشكل الآتي :

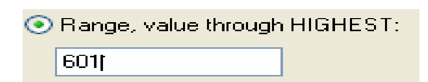


- ثم ننقر كلمة Add فتنتقل القيم المختارة إلى الصندوق Old – New الفئات التالية حتى الفئة قبل الأخيرة هي فئة لها حد أدنى وحد أعلى معرف "151-300" والحد الادني هو 151 والحد الأعلى هو 300 وهكذا باقي الفئات ففي هذه الحالة نقوم فإدخال الحد الأدنى في الخانة العلوية والحد الأعلى في الخانة السفلية كما يلي:



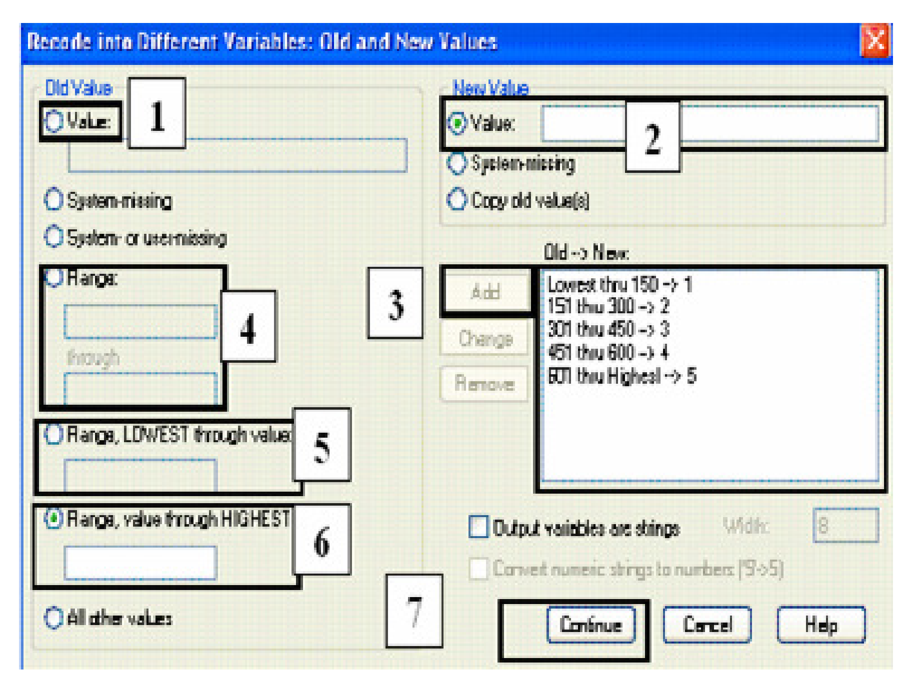
ثم يلي ذلك أن ندخل الترميز "2" في المستطيل الموجود تحت كلمة New value ثم النقر علي add لانتقال إلي إدخال فئة جديدة لتكويد.

- الفئة الأخيرة لها حد ادنى وليس لها حد أعلى وتسمي الفئة المفتوحة من أسفل فنقوم بإدخال الحد الأدنى لها كما يلي:



وبعد ذلك ثم يلي ذلك أن ندخل الترميز "5" في المستطيل الموجود تحت كلمة New value

- عند الانتهاء من عملية التكويد يظهر الشكل (4-11) التالي، في النهاية ثم نختار الأمر Continue لنعود إلى صندوق الحوار الأصلي.



الشكل (4-11)

- نتقر Change ليتم التغيير من قيم المتغير إلى الترميز الجديد.
- ننقر على الأمر Ok فيتم تغيير القيم الأصلية إلى الترميز الجديد حسب الفئات تحت اسم متغير جديد يرمز له بالاسم Newsalary كالآتي:

Salary	Newsalary
100.00	1.00
150.00	1.00
200.00	2.00
250.00	2.00
300.00	2.00
500.00	4.00
270.00	2.00
300.00	2.00
350.00	3.00
700.00	5.00
350.00	3.00
400.00	3.00
550.00	4.00
850.00	5.00

(3-4) تكويد متغير لفظي إلى متغير رقمي:

- ⌖ إنشاء متغيرات وهمية و فئوية باستخدام التكويد

يعتبر التكويد أداة هامة في تحليل العلوم الاجتماعية حيث تكون عملية المقارنة بين المجموعات المختلفة عملية هامة في التحليل مثل المقارنة بين كفاءة الذكور و الإناث في رياضة معينة، المقارنة بين جودة التعليم العامة والخاصة في الجامعات ... الخ. وعلى ذلك فإنه ينبغي لتحليل تلك المقارنات استخدام متغيرات وهمية و متغيرات فئوية.

- ⌖ **المتغيرات الوهمية والمتغيرات الفئوية:**

المتغيرات الوهمية هي المتغيرات التي تأخذ قيمتين عادة ما تكون صفر و واحد. كل قيمة ترمز إلى فئة معينة والقيمة الأخرى ترمز إلى الفئة الثانية.

جدول (4-4)

Value	Category
0	Male
1	Female

- ⌖ المتغيرات الفئوية هي المتغيرات التي تحتوي على أكثر من قيمتين كل قيمة ترمز إلى فئة معينة ولا توجد أفضلية لفئة عن الأخرى مثل أصول الجنسية في الولايات المتحدة تنقسم إلى 6 فئات والتحويل من الفئات إلى القيمة موضح في الجدول التالي:

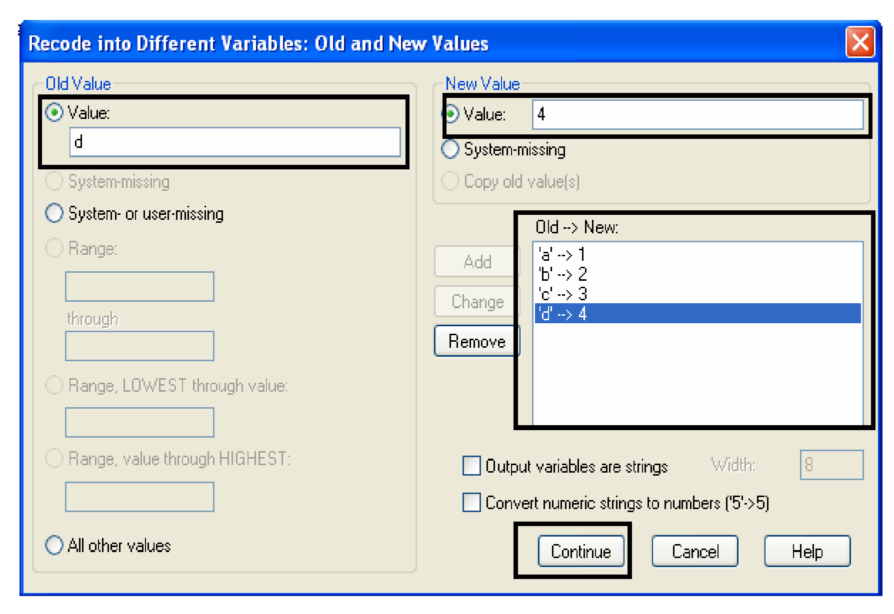
جدول (5-4)

Value	Category
0	White-American
1	African-American
2	Asian-American
3	Hispanic-American
4	Native-American
5	Other

و الآن نقوم بشرح كيفية تحويل تلك المتغيرات باستخدام البرنامج. نفترض أننا نريد التكويد إلى متغير جديد وهي متغير موصوف كما يلي:

kind
a
b
c
d
a
b
c
c
c
d
a
a
b
b

لتنفيذ ذلك نمر بكل الخطوات السابقة إلى أن ينفتح الصندوق الحواري Old And New Values من الأمر Old Values نختار أول مستطيل يسارا من أعلى المسمى Value ويكتب به الحرف المراد تكويده ثم نضع الرقم المناظر لذلك الحرف في أول مستطيل يمينا من أعلى New Value وننقر Add لتضاف إلى القائمة Old-New، نكرر هذه العملية لنحصل على الصندوق الحواري الشكل (4-12):



الشكل(4-12)

ننقر Continue ونمر بنفس الخطوات السابقة إلى أن يتم التغيير فيصبح المتغير بعد التكويد بالشكل الآتي:

kind	newkind
a	1.00
b	2.00
c	3.00
d	4.00
a	1.00
b	2.00
c	3.00
c	3.00
c	3.00
d	4.00
a	1.00
a	1.00
b	2.00
b	2.00

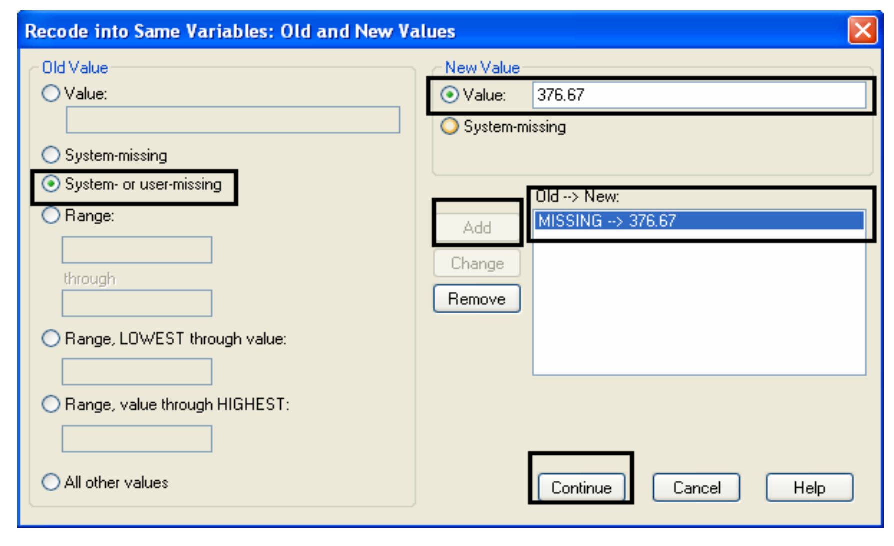
(4-4) إحلال قيم مفقودة بقيم أخرى Replacing Missing Values :

إن وجود قيم مفقودة لبعض المتغيرات تعتبر أحيانا عقبة كبيرة تواجه تطبيق أسلوب إحصائي معين ويتوجب في هذه الحالة تقدير القيم المفقودة حيث يوفر البرنامج SPSS هذه الطريقة. بفرض أن متغير الدخل به مجموعة من القيم المفقودة، تتم عملية الاستبدال عن طريق البرنامج كما يلي:

من القائمة Transform انقر الأمر Recode into same variables فينفتح الصندوق الحواري. اختر المتغير (أو المتغيرات)من قائمة المتغيرات وانقلها إلى المستطيل Variables.

- انقر على Old And New Values لفتح الصندوق الحواري التالي الخاص بعمليات التكويد والمسمى Recode into same Values : Old And New Values.
- في المربع Old Value نختار System-Or User –Missing.
- في المستطيل New Value اكتب القيمة المراد إحلالها مكان القيمة المفقودة ولتكن الوسط الحسابي للبيانات مثلا. ننقر الأمر Add فيتم الاختيار.
- انقر الأمر Continue للعودة للصندوق الأصلي.

انظر الصندوق الآتي بفرض أن الوسط الحسابي 376.67 الشكل (4-13).



الشكل (4-13)

- نتقر الأمر Ok لتنفيذ العملية.

(5-4) الترتيب Ranking:

يمكن بواسطة هذا الأمر تكوين متغيرات جديدة هي عبارة عن رتب متغيرات معينة وتكون هذه الرتب تصاعدية أو تنازلية ويمكن اعطاء رتب لمتغير بواسطة متغيرات أخرى ، حيث أننا في كثير من طرق العرض والتحليل الإحصائي نرغب في التعامل مع الرتب الخاصة بالمتغير وليس بقيم المتغير.

مثال (4-3):

لنفترض أن لدينا المتغيرات التالية: الدخل، الاستهلاك والادخار كالتالي:

جدول (4-6)

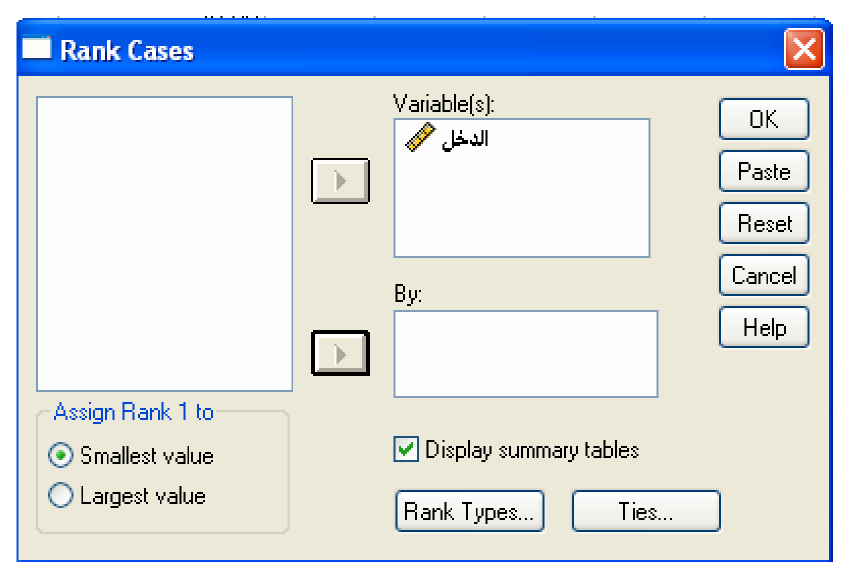
الاسم	الدخل
احمد	50.00
محمود	100.00
مني	150.00
هبة	150.00
امال	200.00
محمد	300.00
شيماء	400.00
عز الدين	500.00
نورا	500.00
عبد التواب	600.00
عمرو	630.00
مهدي	830.00

ونريد إعطاء رتبا تصاعدية لمتغير الدخل .

الحل:

لتنفيذ ذلك نتبع الخطوات التالية:

- من القائمة Transform نختار Rank Cases فيظهر صندوق حوار Rank Cases الذي نرتبه كما في الشكل (4-14):

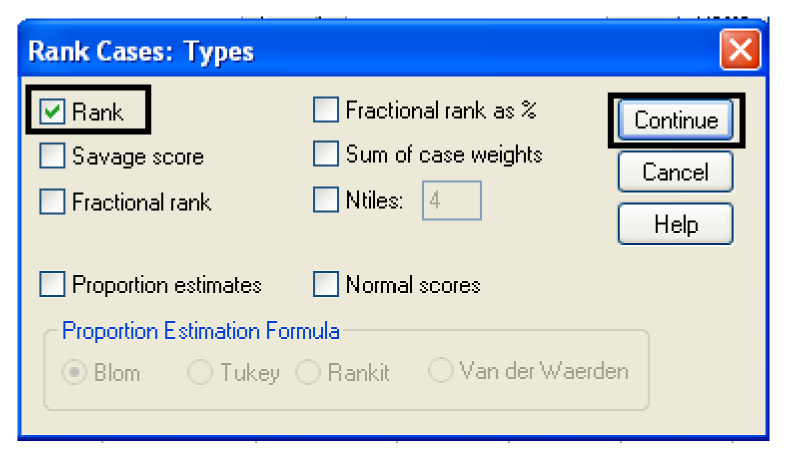


الشكل (4-14)

- لقد قمنا بإدخال العمود الذي نريد إعطاؤه رتبة الدخل في خانة Variables
- وفي خانة Assign Rank To أشرنا الخيار Smallest Value لإعطاء الرتب تصاعديا.
- لاختيار نوع الرتبة انقر زر Rank Types في صندوق حوار Rank Cases يظهر صندوق حوار Rank Cases: Types فيظهر أنواع مختلفة من الرتب:
- Rank: وهي الرتبة البسيطة (حيث يتم إعطاء رتبة لكل قيمة من قيم المتغير تعبر عن ترتيبه ضمن المجموعة).
- Savage Scores: تعطي رتب لقيم المتغير بموجب التوزيع الآسي.
- Fractional Rank: وهي الرتبة الناتجة من قسمة الرتبة البسيطة لقيم المتغير على مجموع الأوزان لكافة الحالات (أو عدد الحالات في حالة عدم وجود أوزان أي يعتبر الوزن مساوياً الي الواحد).
- Fractional Rank as %: يتم الحصول على هذة الرتبة من حاصل ضرب الرتبة السابقة في 100.
- Sum of cases weights : الرتبة تكون متساوية لكافة الحالات وتمثل مجموع الأوزان لكافة الحالات (أو عدد الحالات في حالة عدم وجود أوزان).

⌖ Ntiles: يتم إعطاء رتب بعد تقسيم قيم المتغير الى مجاميع تعطي كل منها رتبة معينة (بعد ترتيبها تصاعدياً أو تنازلياً) فإذا اخترنا 4Ntiles (حالة الترتيب تصاعدي) فإن سيتم إعطاء الرتبة 1 للقيم التي ترتيبها أقل من 25% وتعطي الرتبة 2 للقيم من 25% إلى 50% والرتبة 3 للقيم 50% إلى 75% والرتبة 4 لـ75% فما فوق.

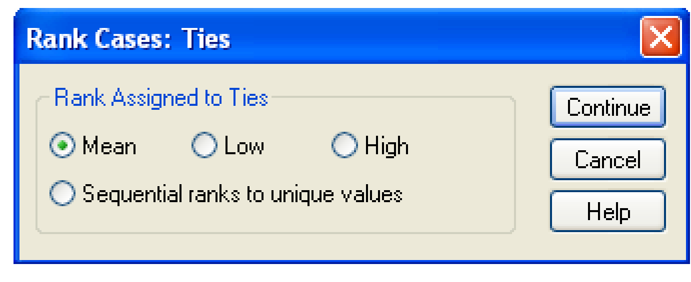
⌖ وهنا في هذا المثال يتم اختيار النوع البسيط للترتيب Rank انظر الشكل(4-15):



الشكل(4-15)

⌖ ثم ننقر على Continue

النقر على الأمر Ties يعطي الصندوق الحواري الفرعي Rank Cases :Ties أنظر الشكل (4-16):



الشكل (4-16)

وفيه يختار المستخدم طريقة التعامل مع التداخلات Rank Assign to Ties هل باستخدام الوسط أم القيمة الصغرى أم الكبرى للرتب المتداخلة.

⌖ ثم زر Ok.

فيضاف متغير جديد (متغير الرتب) باسم الدخل R إلىData Editor كما في الجدول التالي:

جدول (4-7)

الأسم	الدخل	الدخلR
احمد	50.00	1.000
محمود	100.00	2.000
منى	150.00	3.500
هبة	150.00	3.500
امال	200.00	5.000
محمد	300.00	6.000
شيماء	400.00	7.000
عز الدين	500.00	8.500
نورا	500.00	8.500
عبد التواب	600.00	10.000
عمرو	630.00	11.000
مهدي	830.00	12.000

(4-6) **إنشاء سلسلة زمنية** Create Time Series:

هي عبارة عن قيم متغير معين خلال فترات زمنية متساوية كالأيام أو الأشهر أو السنوات، ويمكن اجراء بعض العمليات الإحصائية على السلسلة الزمنية من خلال عدة دوال إحصائية تضمن:الفروق Differences، الأوساط المتحركة Moving Averages، التأخر Lag، التقدم Lead، وغيرها.

مثال (4-4):

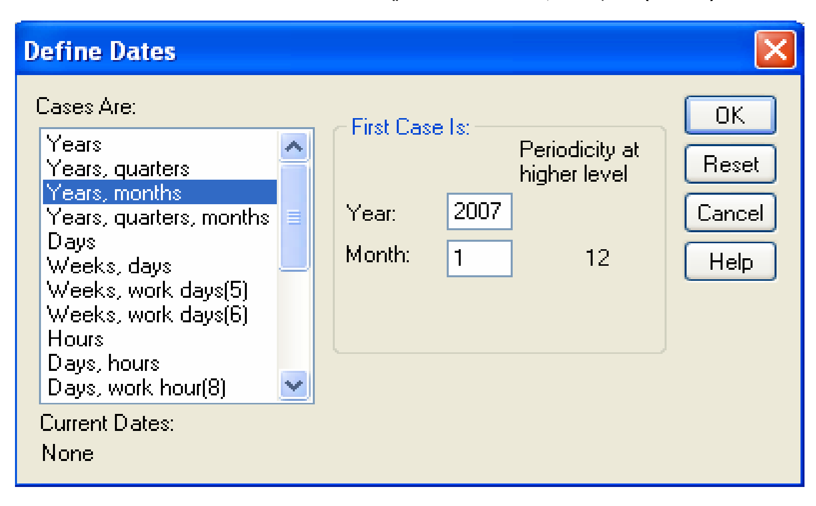
لدينا المتغير حجم المبيعات ويمثل المبيعات الشهرية من أجهزة المحمول خلال 12 شهراً في مؤسسة معينة لسنة 2007 م ونرغب في عمل فروق Differences من الدرجة الأولى لهذا المتغير:

455 445 409 380 268 351 383 382 350 255 207 274

الحل:

- أولاً نقوم بإنشاء السلسلة الزمنية بإنشاء الشهور بالنسبة لسنة 2007 كالتالي:

انقر على الأمر Define Dates من القائمة Data فيظهر صندوق الحوار Define dates الشكل(17-4)الذي نقوم بترتيبه كما يلي:



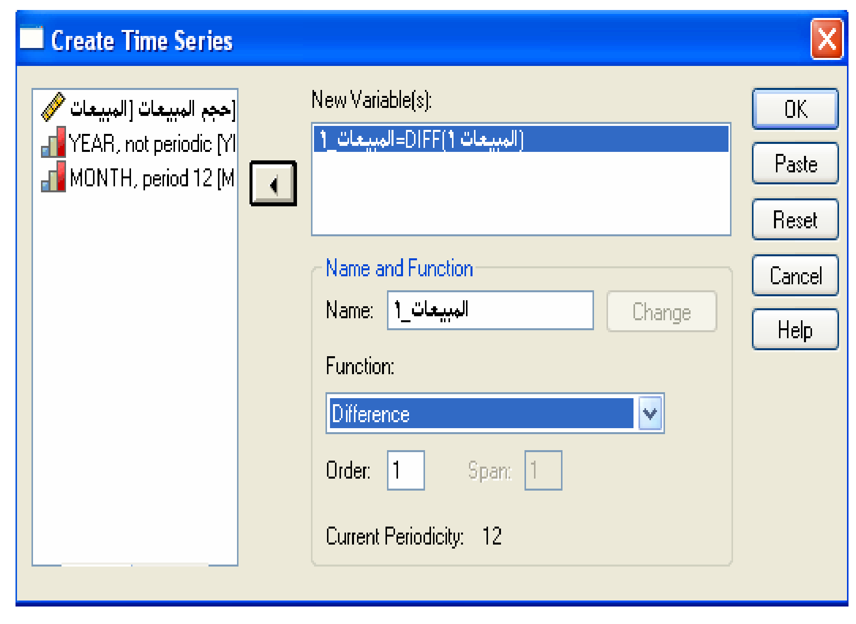
الشكل (17-4)

حيث اخترنا من Cases are السنة والشهور Years،Months ويمكن للباحث أن يختار ما يشاء سواء الأيام أو الساعات أو الأسابيع وغيرها.

- في قائمة First case is نقوم بتعريف التاريخ للحالة الأولى في السلسلة الزمنية: في حقل السنة ندخل سنة البداية 2007، أما في حقل الشهر ندخل شهر البداية 1.
- أماPeriodicity at Higher Level فيبين دورية التاريخ حيث يبين أكبر قيمة يمكن تزويدها للبرنامج مثلاً في المثال أعلى دورية للشهور 12 شهراً.
- عند نقر Ok في صندوق حوار Define Dates تضاف متغيرات التاريخ إلى Data Editor كما يلي:

المبيعات	YEAR_	MONTH_	DATE_
274.00	2007	1	JAN 2007
207.00	2007	2	FEB 2007
255.00	2007	3	MAR 2007
350.00	2007	4	APR 2007
382.00	2007	5	MAY 2007
383.00	2007	6	JUN 2007
351.00	2007	7	JUL 2007
268.00	2007	8	AUG 2007
380.00	2007	9	SEP 2007
409.00	2007	10	OCT 2007
445.00	2007	11	NOV 2007
455.00	2007	12	DEC 2007

- ثانياً نقوم بعمل الفروقات Differences بأتباع الخطوات التالية :
- انقر على الأمر Create Time Series من القائمة Transform، يظهر الصندوق الحواري Create Time Series حدد اسم المتغير الذي تعتمد عليه السلسلة وانقله إلى المستطيل New Variables أنظر الشكل (4-18).



الشكل (4-18)

حدد اسم المتغير الجديد في المستطيل Name تحت الأمر Name and Function وليكن "المبيعات _1". يتم تحديد الدالة الجديدة التي سيتم على أساسها إنشاء بيانات

المتغير الجديد من الأمر Function والتي في المثال هي Differences، المتغير الجديد المختار هو دالة الفروق.

عند نقر Ok يضاف متغير جديد باسم المبيعات _1 إلى Data Editor كم يلي في الشكل (4-19):

المبيعات	YEAR_	MONTH_	DATE_	المبيعات_1
274.00	2007	1	JAN 2007	.
207.00	2007	2	FEB 2007	-67.00
255.00	2007	3	MAR 2007	48.00
350.00	2007	4	APR 2007	95.00
382.00	2007	5	MAY 2007	32.00
383.00	2007	6	JUN 2007	1.00
351.00	2007	7	JUL 2007	-32.00
268.00	2007	8	AUG 2007	-83.00
380.00	2007	9	SEP 2007	112.00
409.00	2007	10	OCT 2007	29.00
445.00	2007	11	NOV 2007	36.00
455.00	2007	12	DEC 2007	10.00

الشكل (4-19)

الفصل الخامس

تحليل متغير واحد

Univariate Analysis

(1-5) مقدمة:

يبدأ التحليل الإحصائي بتحليل السمات الإحصائية لكل متغير على حدة والذي نطلق عليه التحليل الأحادي Univariate Analysis ومن خلال كل تحليل نقوم بمعرفة الآتي:

- كيف تتوزع قيم المفردات- توزيع طبيعي، أسي، ذي الحدين...،
- معرفة مقاييس النزعة المركزية وهي مجموعة من المقاييس تصف المتغير مثل الوسط ، الوسيط ، المنوال...،
- معرفة تشتت المفردات عن القيم المركزية المحسوبة في الخطوة السابقة وتقاس عادة باستخدام التباين، الانحراف المعياري، المدى ...،
- توضيح هل توجد قيم شاذة في المفردات من خلال مقياس معامل الاختلاف.
- إجراء اختبارات الفروض على البيانات إذا كانت البيانات بيانات عينة عشوائية مأخوذة من مجتمع إحصائي.

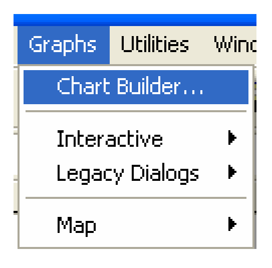
في هذا الفصل نقوم بتوضيح الطرق المختلفة لعرض وتحليل متغير أحادي ففي المبحث الأول نقوم بتوضيح كيف يمكن عرض البيانات في أشكال بيانية أو مخططات رسومية للمتغير. ونستكمل فيما تبقى من الفصل كيفية إيجاد المقاييس الإحصائية المناسبة للمتغير.

(2-5) عرض البيانات باستخدام

الرسومات البيانية البسيطة Simple Bar Graphs

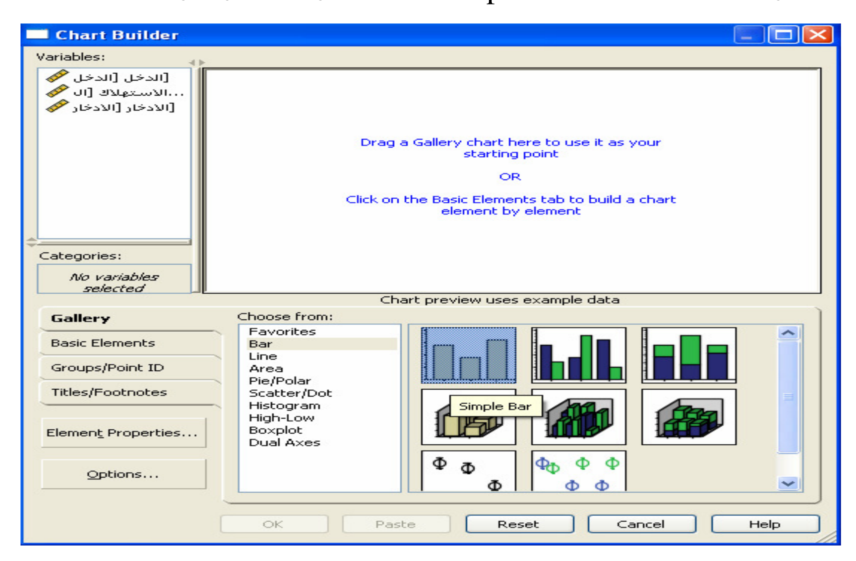
الرسومات البيانية نستطيع أن نستخدمها للحصول على معلومات محددة عن المتغير مثل الوسط، الوسيط، التوزيع المتراكم ولعمل رسم بياني نتبع الخطوات التالية:

- نختار Graphs ثم نقوم باختيار ...Chart Builder من الشكل الآتي:



شكل (1-5)

- فيظهر الشكل (2-5) فنختار منه Bar من عمود Choose from: ثم نقوم باختيار الشكل الأول وهو Simple Bar وذلك بالنقر عليه نقراً مزدوجا:



شكل (2-5)

- بعد النقر على الشكل نقراً مزدوجا يظهر الشكل (3-5) الآتي والذي يتكون من جزئين بالشكل الآتي:

أ- الجزء الأيمن هو شكل الرسم البياني والمتغيرات المتاحة للرسم.

ب- الجزء الايسر و هو شكل خصائص المحاور المتاحة.

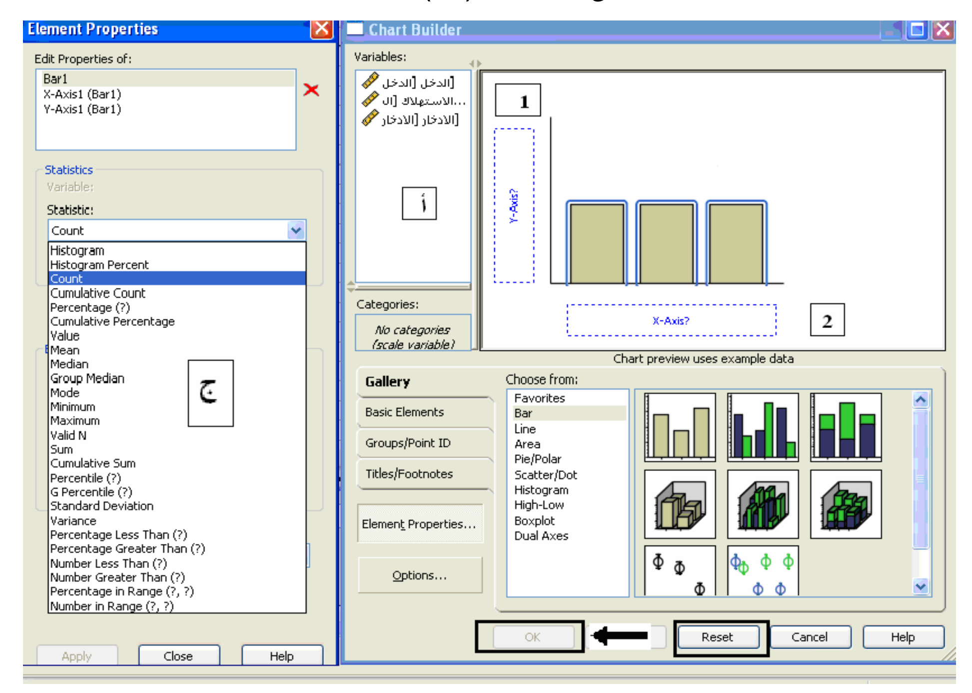
و هنا يوجد داخل شاشة الرسم ما يلي:

- "أ" وهو عمود المتغيرات و هنا نفترض أننا نريد رسم المتغير الدخل عن بيانات المثال (3-4) -انظر شكل (3-4).

- "2" وهو المحور X في الشكل البياني المراد رسمه ونفترض أننا نريد رسم المتغير الدخل على المحور Xفإننا ننقر عليه ونسحبه إلي محور X والمحدد في الشكل برقم 2.

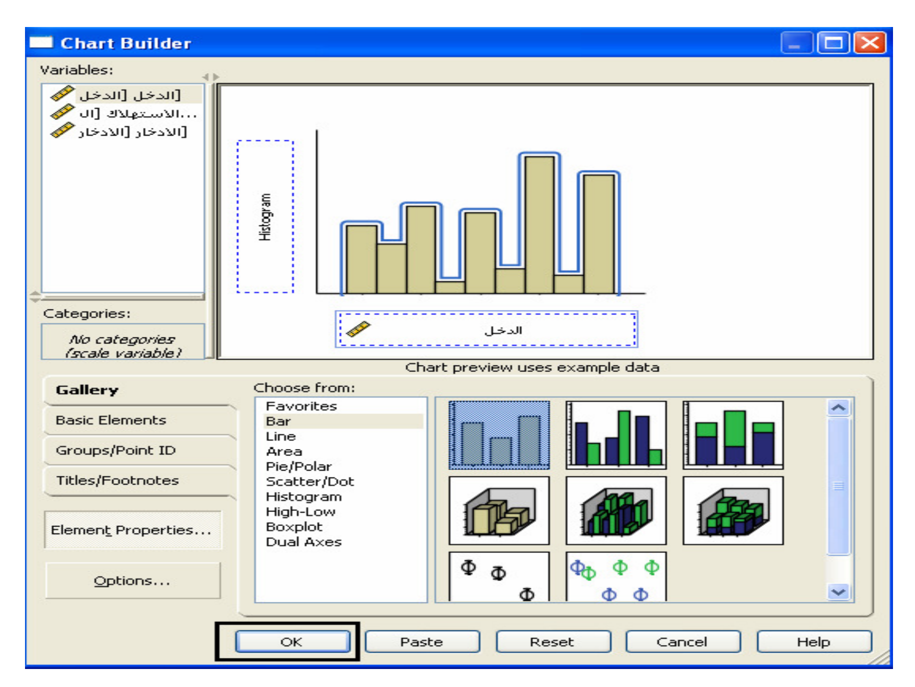
- "1" Y-Axis? وهو محور العمليات التي نوضح عليها الشكل و هنا نفترض أننا نريد أن نوضح الشكل التكراري للمتغير Histogram فإننا ننتقل إلى الجزء الثاني من الشكل (3-5).

⌖ "ج" وهو عمود يحمل اسم Statistics وهو عمود العمليات التي نريد اظهارها على المحور Y ونستطيع الاختيار المتعدد لها ولنفرض أننا نريد اظهار تكرار قيم متغير الدخل فإننا نختار Histogram من (جـ).

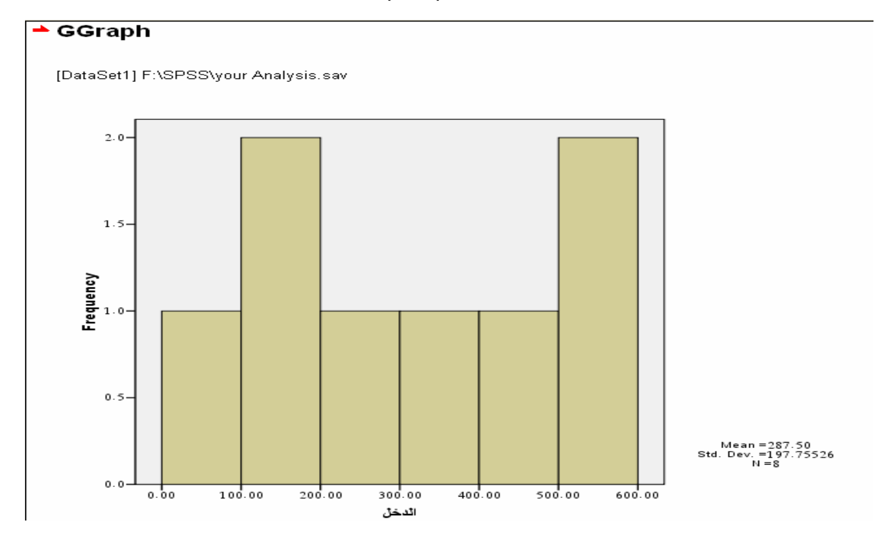


شكل (3-5)

⌖ بعد الإنتهاء من العمليات السابقة يصبح الشكل السابق على الشكل التالي ومن ثم النقر على OK.



شكل (4-5)

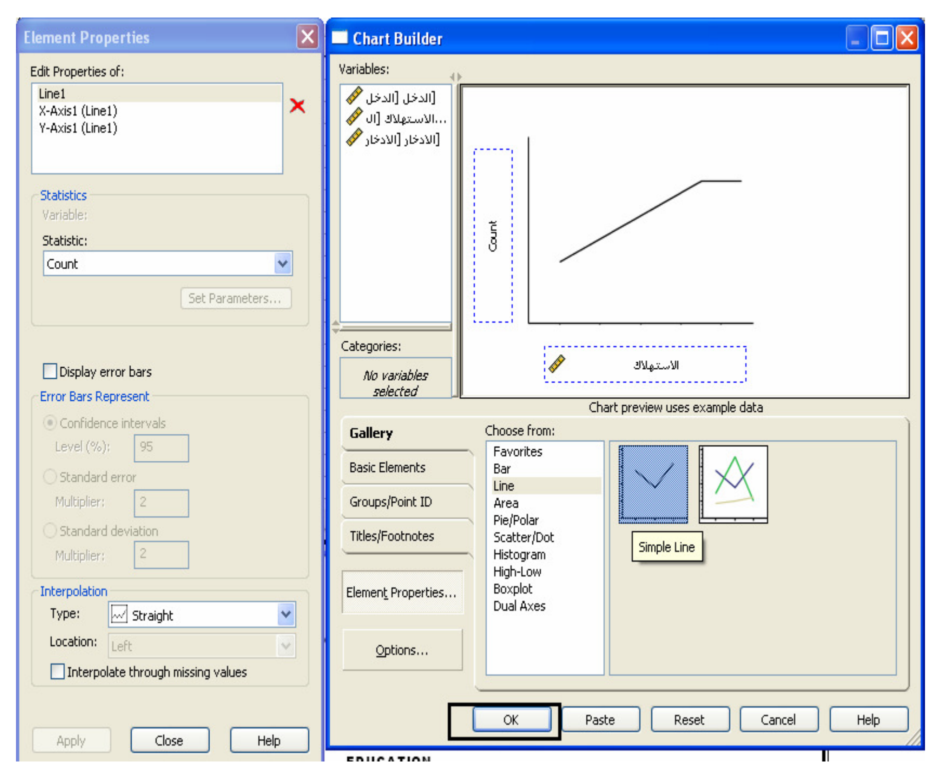


شكل (5-5)

- ونلاحظ إن الشكل السابق يوضح الرسم البياني لمتغير الدخل وبجوار الرسم بعض المقاييس الإحصائية عن المتغير مثل الوسط الحسابي mean يساوي 287.5، الانحراف المعياري يساوي 197.75526 وعدد البيانات يساوي 8.

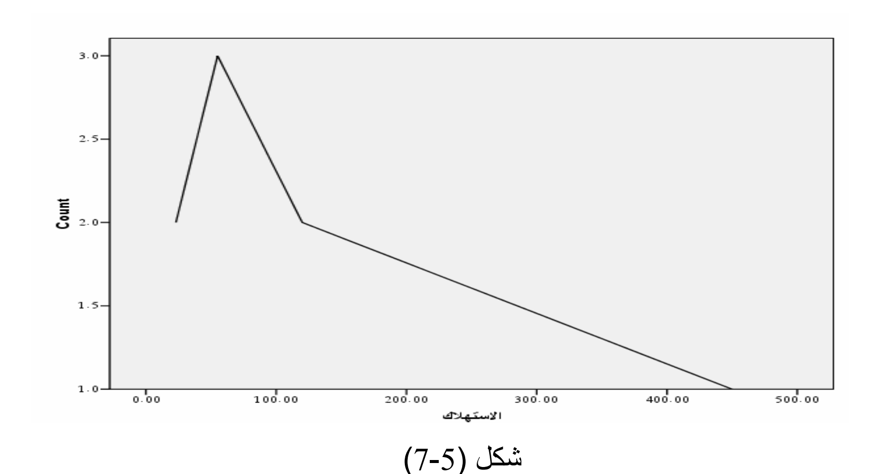
⌖ Line Graphs

إذا أردنـا عمـل شـكل بيـاني مـن نـوع line للمتغيـر الاسـتهلاك فإننـا نختـار Graphs ثم نقوم باختيار ...Chart Builder فيظهر الشكل (5-1) فنختار منه Line من عمود :Choose from ثم نقوم باختيـار الشكل الأول وهو Simple Line وذلك بالنقر عليه نقراً مزدوجا ومن ثم اختيار المتغير الاستهلاك من عمود المتغيرات بنفس طريقة اختيار الدخل السابق وذلك بسحبها إلى عمود X فيظهر الشكل (5-6) الآتي:



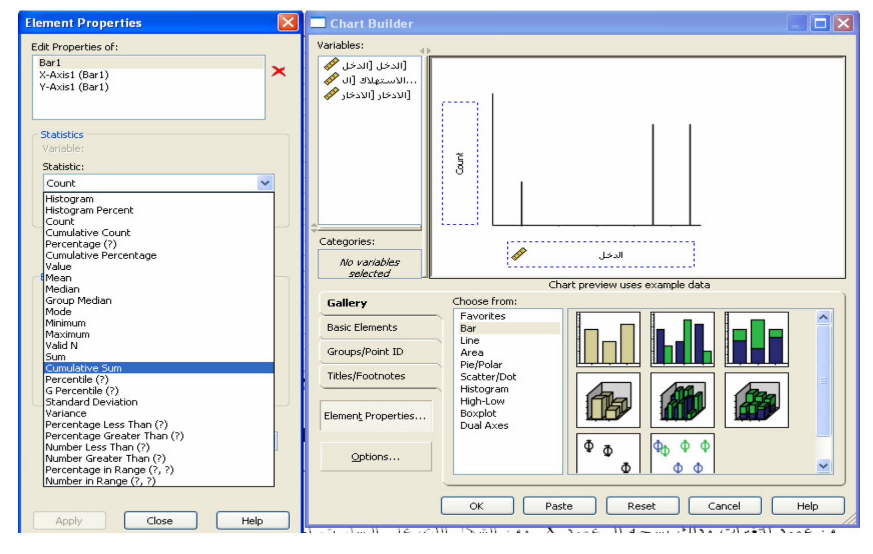
شكل (5-6)

⌖ وبعد النقر علي ok لتأكيد الاختيـار فإننـا الشـكل (5-7) الآتـي يظهر فـي نافذة output.



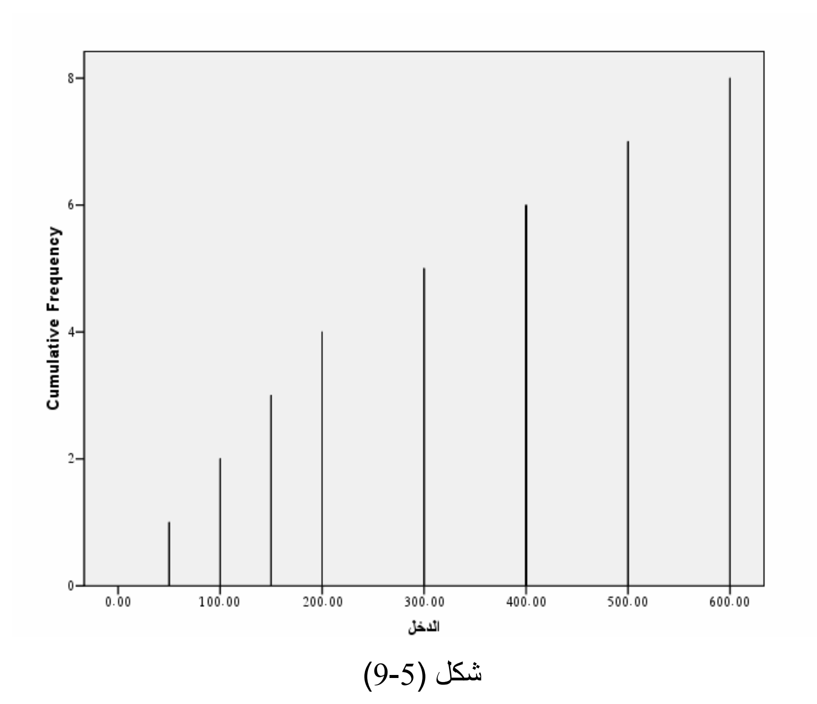
شكل (5-7)

إذا أردنا عمل شكل بياني من نوع Graphs for cumulative frequency للمتغير الدخل فإننا نختار Graphs ثم نقوم باختيار Chart Builder… فيظهر الشكل (5-1) فنختار منه Bar من عمود Choose from: ثم نقوم باختيار الشكل الأول وهو Simple Bar وذلك بالنقر عليه نقراً مزدوجا ومن ثم اختيار المتغير الدخل من عمود المتغيرات وذلك بسحبه إلى عمود X ومن الجزء الذي على اليسار يتم الاختيار فيظهر الشكل (5-8) الآتي واختيار cumulative sum من عمود Statistic:



شكل (5-8)

وبعد ذلك النقر على Ok فيظهر الشكل الآتي (5-9) الآتي :



شكل (5-9)

⌖ الحصول على مقياس إحصائي Statistical Measure:

في كثير من النواحي التطبيقية يكون الباحث في حاجة إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تنزع إليها القيم وتسمي مقاييس النزعة المركزية، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير وتسمي مقاييس التشتت، وأيضا ما إذا كان هناك قيم شاذة أم لا من خلال مقياس يسمي معامل الاختلاف. والاعتماد على العرض البياني وحدة لا يكفى، ولذا يتناول هذا الفصل كيف يمكن وصف المتغير باستخدام المقاييس الإحصائية المناسبة، توجد مجموعة رئيسية من المقاييس تعبر عن البيان الإحصائي منها:

(3-5) مقاييس النزعة المركزية Measures Of Central Tendency :

⌖ **الوسط الحسابي** Arithmetic Mean:

يعتبر الوسط الحسابي من أهم المقاييس في التحليل الإحصائي بشكل عام إلا أنه يناسب فقط الظواهر الكمية فقط ونستطيع تعريفه على أنه مجموع قيم المشاهدات على عددها، وتكتب وتحسب بشكل رياضي من خلال العلاقة التالية:

$$\overline{X} = \sum_{i=1}^{n} x_i / n , \quad \mathrm{i} = 1, 2 \ldots \mathrm{n}$$

حيث:

$\overline{X}$ رمز للمتوسط وتقرأ "اكس بار"،

$\sum_{i=1}^{n}$ رمز للمجموع وتقرأ "سيجما من i=1 إلى n"،

n عدد المفردات محل الدراسة.

- **الوسيط** Median:

الوسيط هو أحد أهم مقاييس النزعة المركزية ويمكن أن تكون الظـاهرة أو المتغير كمـي أو ترتيبـي، ويعـرف بأنـه القيمـة التـي تقـع فـي منتصـف البيانـات بعـد الترتيـب (تصاعديا أو تنازليا)

وهنا توجد حالتان بالنسبة للبيانات الكمية:

- أن تكـون عـدد البيانـات فرديـاً وفـي هـذه الحالـة الوسـيط هـو القـيم التـي عنـد الترتيـب $\frac{n+1}{2}$ حيث أن n هي عدد المفردات وفي هذه الحالة نرمز للوسيط ونحسبه باستخدام العلاقة الآتي:

$$median = X_{\frac{n+1}{2}}$$

- أن تكون عدد البيانـات زوجيـاً وفـي هـذه الحالـة الوسيط هـو الوسـط الحسـابي للقيمتين التي عند الترتيب $\frac{n}{2}$ و $\frac{n}{2}+1$ وفي هذه الحالة نرمز للوسيط ونحسبه باستخدام العلاقة الآتي:

$$median = \frac{\left[X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right]}{2}$$

- **المنوال** Mode:

وهو القيمة الأكثر تكرارا في البيانات وقد يكون للبيانات أكثر من منوال أو قد ينعدم أي لا وجود له.

مقاييس التشتت المطلق والنسبي:

هو مقياس لدرجة التجانس أو تشتت مفردات البيانات بعضها عن بعض .

- **المدى** Range:

ويعرف بأنه الفرق بين أكبر رقم وأصغر رقم للبيانات وهو كالتالي:

المدى R = أكبر قراء – أصغر قراءة.

- **الانحراف المعياري** Standard Deviation (σ):

يقيس تشتت البيانات حول المتوسط وهو الجذر التربيعي الموجب للتباين σ^2 ويحسب كالآتي:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n}}$$

حيث:

σ رمز للانحراف المعياري وتقرأ "سيجما"

σ^2 رمز التباين وتقرأ "سيجما تربيع"

- **معامل الاختلاف** Coefficient Of Variation γ :

وهو من المقاييس النسبية للتشتت ويعرف أنه خارج قسمة الانحراف المعياري على الوسط أي

$$\gamma = \frac{\sigma}{\bar{x}}$$

وهو من أفضل مقاييس التشتت النسبية، ويستخدم للتعرف على وجود قيم شاذة للبيانات أم لا.

- **مقياس الالتواء** Measure Of Skewness α :

وهو مقياس يحدد هل البيانات متماثلة أم ملتوية، يعرف التماثل بأنه إذا رسمنا من قمة المنحنى التكراري للبيانات عمود لقسم المساحة تحت المنحنى إلى جزئيين متماثلين، ويعرف الالتواء بالبعد عن التماثل ، الالتواء قد يكون بذيل لليمين أو بذيل لليسار.

- **مقياس التفرطح** Measure Of Kurtosis

عندما يتركز عدد أكبر من القيم بالقرب من منتصف المنحنى، ويقل في طرفيه،

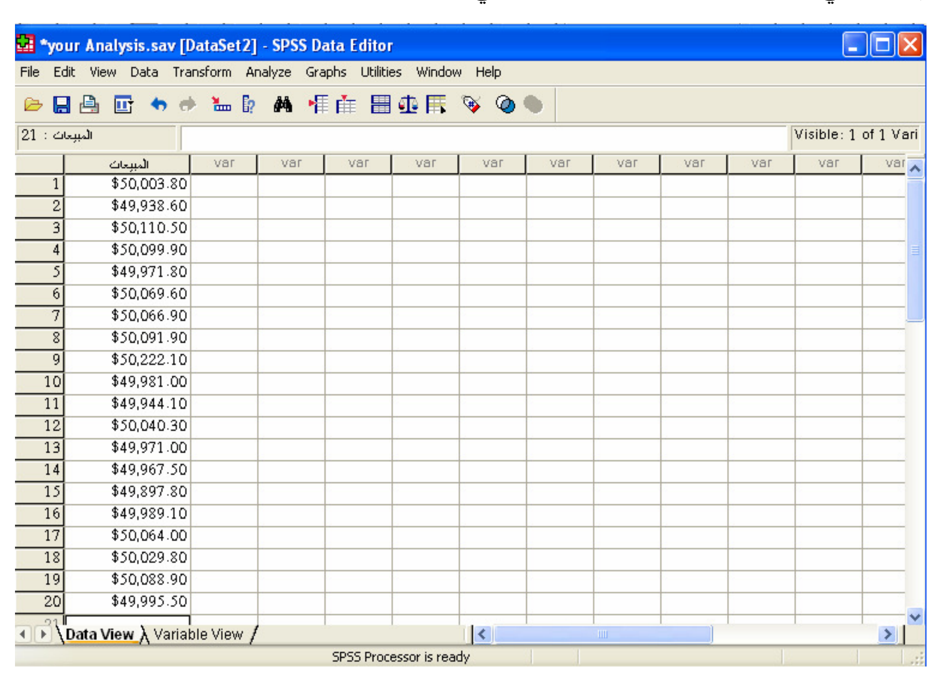
يكون المنحنى مدببا، وعندما يتركز عدد أكبر على طرفي المنحنى ، ويقل بالقرب من المنتصف يكون المنحنى مفرطحا ، أو منبسطا.

الآن نستطيع ايجاد المقاييس الاحصائيه السابقه وغيرها من خلال برنامجنا SPSS وما يحويه من أوامر تمكننا من تنفيذ وتحديد كل المقاييس السابقة فضلا على إيجاد توزيع تكراري للبيانات لمتغير واحد أو أكثر.

⌖ **الأمر Frequencies:**

يستعمل هذا الأمر لعرض تكرار كل قيمة لمتغير ما وحساب بعض مقاييس التشتت والنزعة المركزية و الربيعيات والمئينات مع عرض بعض المخططات البيانية يمكن الحصول على هذا الأمر من القائمة Analyze ثم الأمر Frequencies.

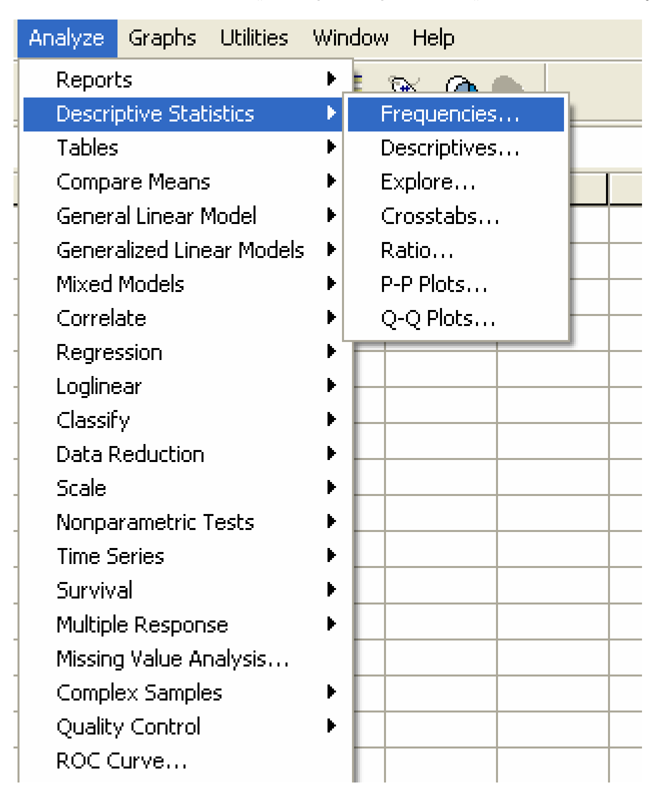
إذا افترضنا أننا لدينا بيانات عن قيمة مبيعات منتج السيارات لشركة ما وتم إدخالها في شاشة data view وكانت كالآتي :



شكل (5-10)

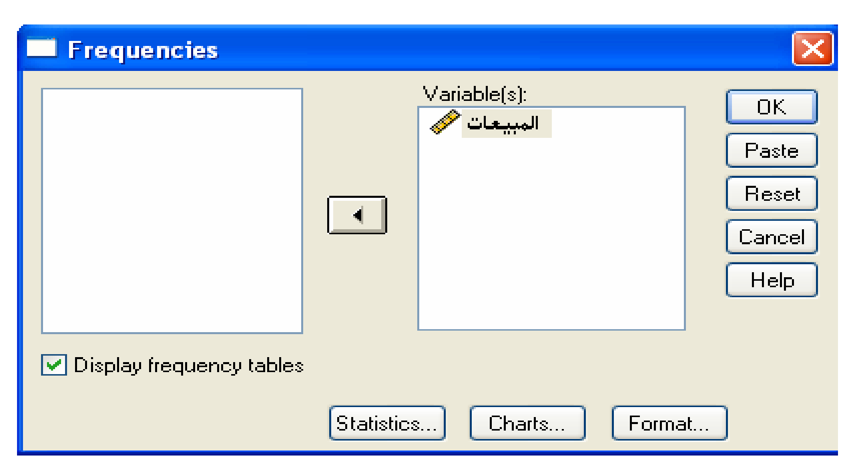
ونريد إيجاد توزيع تكراري له وأيضا إيجاد مقاييس النزعة المركزية ومقاييس التشتت كالتالي:

من القائمة Analyze الظاهر في الشكل (5-11) الآتي:



الشكل (5-11)

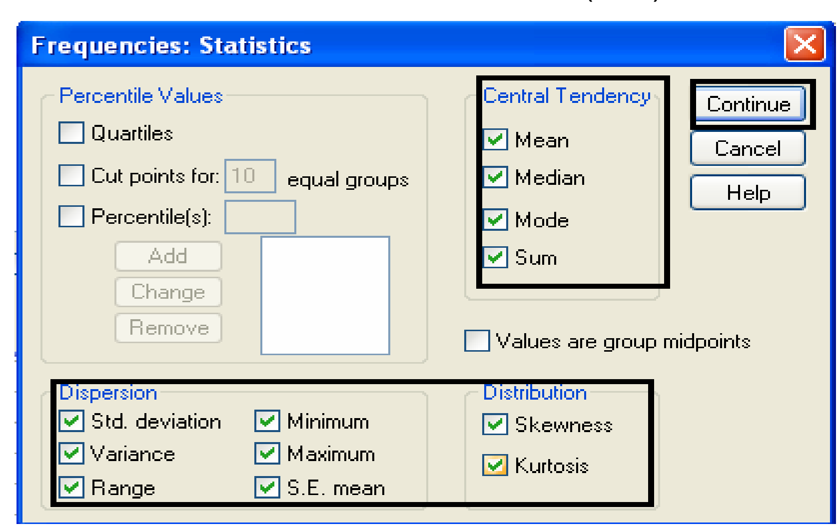
- انقر على Descriptive statistics ثم انقر على الأمر الفرعي Frequencies فيظهر صندوق حوار Frequencies الذي يظهر الشكل التالي بعد إدخال المبيعات عن طريق تحديدها ثم النقر على علامة الانتقال [▶] الذي يرمز للقيم في قائمة Variables انظر الشكل (5-12):



الشكل (5-12)

حيث أن:

- ⌖ - Display Frequency Tables: يعرض الجدول التكراري عند تأشيرة المربع المجاور له.
- ⌖ - Statistics: يعرض بعض المؤشرات الإحصائية حسب حاجة المستفيد ، عند نقره يظهر صندوق حواري Statistics وقد اخترنا المؤشرات التالية للعرض كما يلي انظر الشكل (5-13):

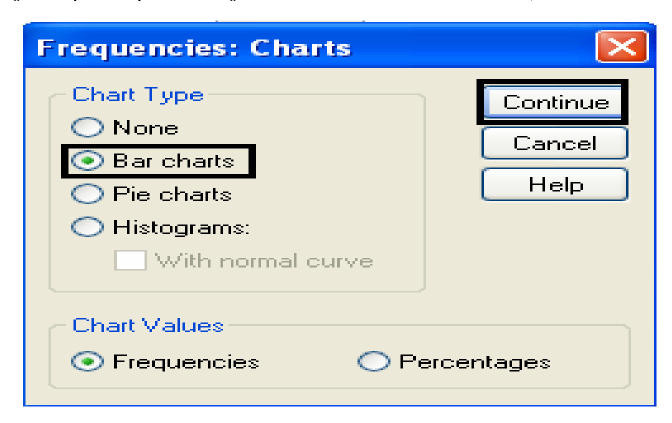


الشكل(5-13)

- Percentile Values: تشـمل الربيعـات Quartile والمئينـات Percentiles ولتحديد المئين نكتب المئين في المربع المجاور لـ Percentiles ثم ننقر زر Add

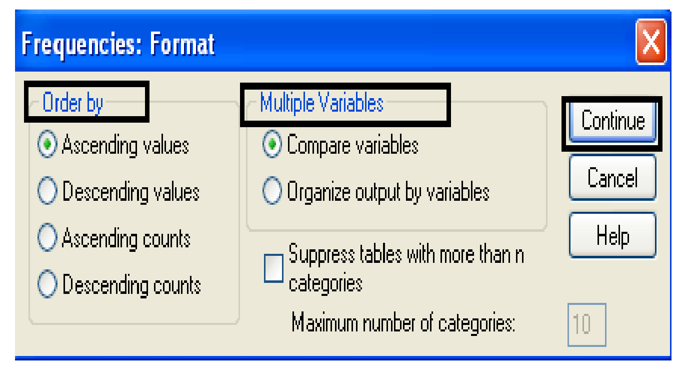
فيتم إضافة المئين الذي نريد إلى المستطيل أسفل Percentiles ويستعمل الزر Change لتغيير قيمة المئين والزر Remove لإزالة المئين.

- Dispersion: لعرض مقاييس التشتت.
- Central Tendency: لعرض مقاييس النزعة المركزية.
- وبعد تحديد ما نريد من مقاييس ننقر على continue .
- Chart: لعرض المخططات Bar ، Pie وغيرها من المخططات ونفرض أننا نريد Bar فإننا نحددها ثم ننقر على continue كما في الشكل (14-5) الآتي:



الشكل (14-5)

- Format: عند النقر عليه يظهر صندوق الحوار أنظر الشكل (15-5):



الشكل (15-5)

حيث أن:-

- Ordered By: لترتيب المشاهدات في الجدول التكرار تصاعديا أو تنازليا حسب القيم Values أو التكرارات Counts وفي هذا المثال اخترنا الترتيب حسب القيم.
- Multiple Variables : يستخدم هذا الأمر في حالة وجود أكثر من متغير في القائمة Variables في صندوق حوار Frequencies وتشمل ما يلي:
- ✓ Compare Variables: لعرض المؤشرات الإحصائية للمتغيرات كافة في جدول واحد.
- ✓ Organize Out By Variables: لعرض مؤشرات كل متغير في جدول مستقل وفي هذا المثال لا يهم اختيار أي من هذين المتغيرين لوجود متغير واحد فقط.
- ✓ Suppress Tables With More Than n Categories: لإخفاء الجدول التكراري للمتغيرات التي يزيد عدد فئاتها عن العدد المحدد من قبل المستفيد ونلاحظ عند تنشيطها يتم تنشيط خانة بعنوان Maximum numbers of categories عندها نستطيع ما هو العدد المحدد الذي نريد إظهاره فقط.
- عند النقر على زر Ok في صندوق حوار Frequencies تظهر النتائج التالية:

جدول (1-5)

Statistics

المبيعات

N	Valid	20
	Missing	0
Mean		########
Std. Error of Mean		$16.97829
Median		########
Mode		########a
Std. Deviation		$75.92924
Variance		5765.250
Skewness		.650
Std. Error of Skewness		.512
Kurtosis		.782
Std. Error of Kurtosis		.992
Range		$324.30
Minimum		########
Maximum		########
Sum		$1000544

a. Multiple modes exist. The smallest value is shown

حيث أن:

- Valid: تمثل القيم الصحيحة (الغير مفقودة).
- Missing: تمثل القيم المفقودة.

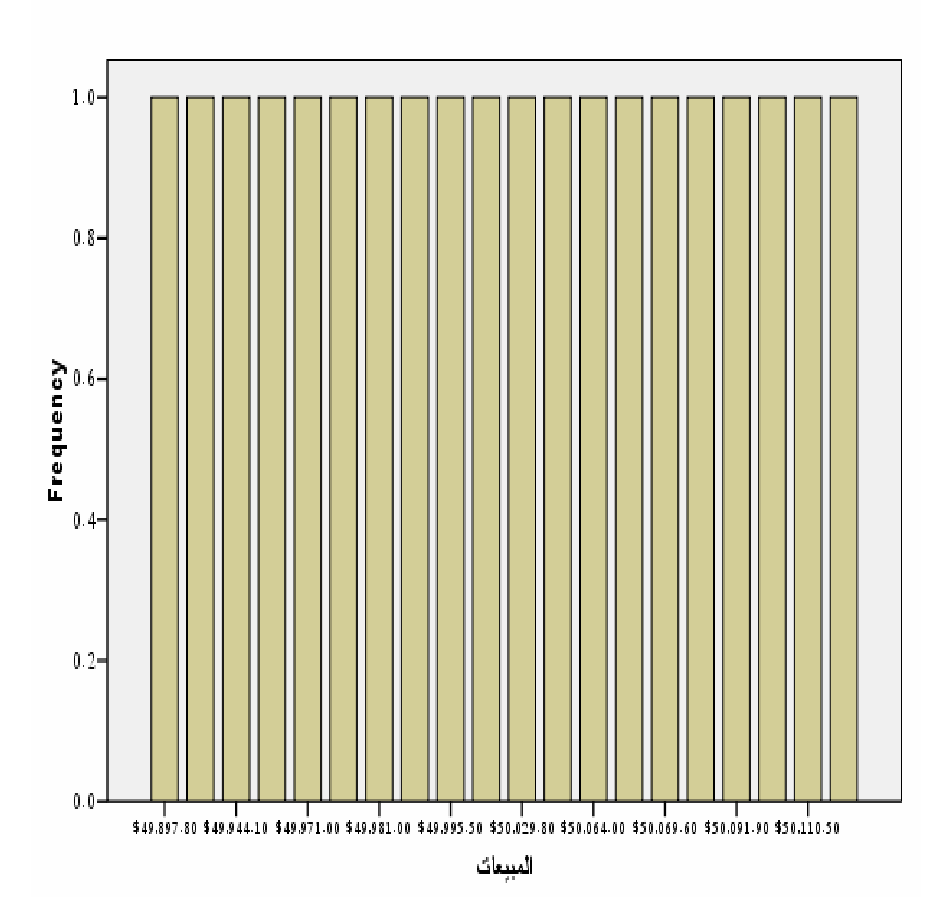
جدول (2-5)

المبيعات

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	$49,897.80	1	5.0	5.0	5.0
	$49,938.60	1	5.0	5.0	10.0
	$49,944.10	1	5.0	5.0	15.0
	$49,967.50	1	5.0	5.0	20.0
	$49,971.00	1	5.0	5.0	25.0
	$49,971.80	1	5.0	5.0	30.0
	$49,981.00	1	5.0	5.0	35.0
	$49,989.10	1	5.0	5.0	40.0
	$49,995.50	1	5.0	5.0	45.0
	$50,003.80	1	5.0	5.0	50.0
	$50,029.80	1	5.0	5.0	55.0
	$50,040.30	1	5.0	5.0	60.0
	$50,064.00	1	5.0	5.0	65.0
	$50,066.90	1	5.0	5.0	70.0
	$50,069.60	1	5.0	5.0	75.0
	$50,088.90	1	5.0	5.0	80.0
	$50,091.90	1	5.0	5.0	85.0
	$50,099.90	1	5.0	5.0	90.0
	$50,110.50	1	5.0	5.0	95.0
	$50,222.10	1	5.0	5.0	100.0
	Total	20	100.0	100.0	

نلاحظ أن المشاهدات في الجدول قد رتبت تصاعديا حسب قيم المتغير X.





ونلاحظ أن كل البيانات لها تكرار واحد ولا يوجد قيمة من البيانات تم تكرارها مرتين أو أكثر.

⌖ **الأمر Descriptive:**

من القائمة Analyze يمكن الحصول على الأمر Descriptive ويمكن أيضا إجراء العرض الإحصائي للبيانات عن طريقه أيضا وذلك بالحصول على المقاييس الإحصائية المختلفة وعلى بعض الأشكال البيانية ،هذا الأمر يسمح أيضا بالحصول على ما يسمى بالقيم المعيارية ويتم حساب القيم المعيارية وفق العلاقة:

$$Z = \frac{X - \mu}{S}$$

حيث X هي القيمة المدخلة ، μ هي المتوسط الحسابي للمتغير X وتقرأ (ميو)، S هو الانحراف المعياري للمتغير X والقيم المعيارية تظهر عادة في نهاية ملف البيانات.

مثال (5-1):

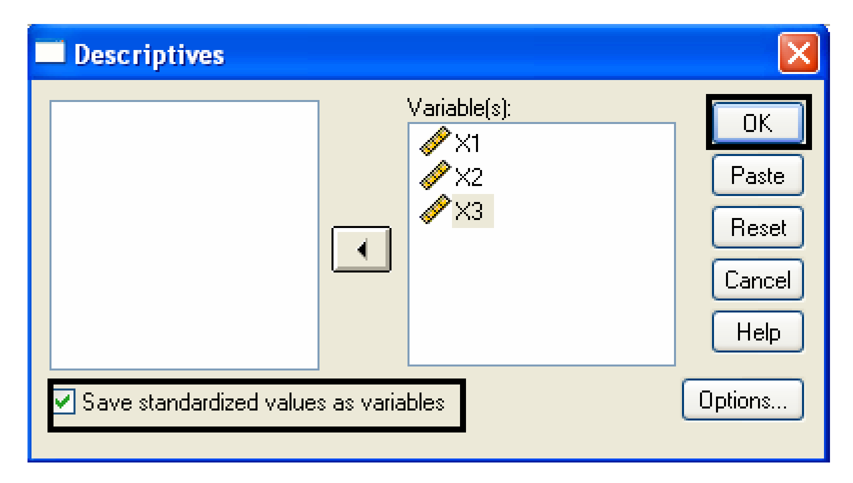
الجدول التالي يتضمن المتغيرات X_1، X_2، X_3 والتي تم إدخالها في Data editor للبرنامج:

جدول (5-3)

X_1	X_2	X_3
90	50	12
70	52	15
56	55	19
65	60	22
85	65	20
60		
69		
57		
50		
75		
62		
51		
85		

لغرض استخراج المقاييس الوصفية بالأمر Descriptive نتبع الخطوات التالية :

- نختار من القائمة Analyze الخيار Descriptive Statistics ومن القائمة الفرعية الخيار Descriptive يظهر مربع الحوار الذي نقوم من خلاله بإدخال المتغيرات على الشكل انظر الشكل (5-16):

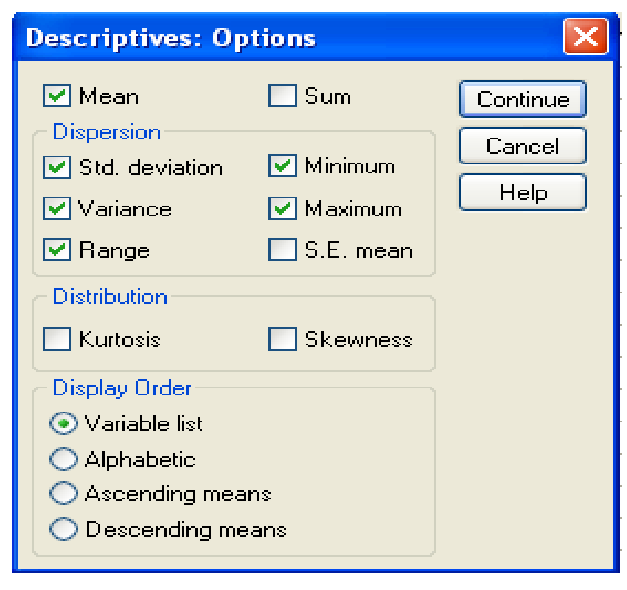


الشكل (5-16)

حيث أن :

- Save Standardized Values As Variables: لإضافة المتغيرات المعيارية z إلى Data editor عند التأشير عليها في مربع الحوار.

Options: عند نقره يظهر صندوق الحوار انظر الشكل (5-17):



الشكل (5-17)

اختر المقاييس المطلوبة.

✓ Display Order :يمكننا من خلالها ترتيب عرض المقاييس الوصفية للمتغيرات، وقد اخترنا Variable List: الذي يعرض المقاييس الوصفية حسب تسلسل المتغيرات الوارد في خانة Variables في صندوق حوار Descriptive.

✓ ثم انقر على Continue لنعود لمربع الحوار Descriptive.

✓ ثم ننقر على زر Ok في صندوق حوار Descriptive يعرض المخرجات التالية:

جدول (5-4)

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
X1	13	40.00	50.00	90.00	67.3077	13.21858	174.731
X2	5	15.00	50.00	65.00	56.4000	6.10737	37.300
X3	5	10.00	12.00	22.00	17.6000	4.03733	16.300
Valid N (listwise)	5						

حيث تم عرض المقاييس الإحصائية المختارة حسب ترتيب المتغيرات في خانة Variables في صندوق حوار Descriptive، كما تم إضافة القيم المعيارية للمتغيرات إلى Data editor كما في الشكل التالي:

X1	X2	X3	ZX1	ZX2	ZX3
90.00	50.00	12.00	1.71670	-1.04791	-1.38706
70.00	52.00	15.00	.20368	-.72044	-.64399
56.00	55.00	19.00	-.85544	-.22923	.34676
65.00	60.00	22.00	-.17458	.58945	1.08983
85.00	65.00	20.00	1.33844	1.40813	.59445
60.00	.	.	-.55284	.	.
69.00	.	.	.12802	.	.
57.00	.	.	-.77979	.	.
50.00	.	.	-1.30935	.	.
75.00	.	.	.58193	.	.
62.00	.	.	-.40153	.	.
51.00	.	.	-1.23370	.	.
85.00	.	.	1.33844	.	.

الشكل (5-18)

الفـصل السـادس

تحليل متغيـريـن

Bivariat Analysis

(1-6) مقدمة:

بعد عرض وصف متغير واحد في الفصل السابق يكون من الضروري هنا أن نوضح كيفية وصف متغيرين أو أكثر حيث نجد أن العلاقة بين متغيرين تكون من أهم الأشياء التي لاغنى عنها في التحليل الإحصائي. ولذلك كان من المهم التعرف على كيفية وصف متغيرين وتوضيح العلاقة بينهما باستخدام أشكال الانتشار والجداول المزدوجة و هذا بالاضافة الى المقاييس الإحصائية المناسبة لهم.

توجد ثلاث طرق للعرض والتلخيص في حالة متغيرين:

- العرض الجدولي .
- العرض البياني.
- العرض باستخدام أحد المقاييس الإحصائية.

ويمكن عرض البيانات المزدوجة كالآتي:

- الجدول المزدوج لمتغيرين : هو جدول ثنائي تعرض فيه بيانات المتغيرين معا وهو يصف العلاقة بين المتغيرين بالإضافة إلى وصف بيانات كل متغير منفردا .
- تمثيل بيانات الجدول المزدوج بيانيا : إذا كانت البيانات كمية فإنه يمكن تمثيل البيانات باستخدام شكل الانتشار وإذا كانت البيانات وصفية فيتم تمثيلها بيانيا على ما يسمى بالأعمدة البيانية المزدوجة.
- حساب معامل الارتباط بين المتغيرين: فيوضح معامل الارتباط مدى قوة العلاقة بين متغيرين فيحدد درجة قوتها (قوية، ضعيفة، تامة، غير موجودة) واتجاهها (عكسية، طردية) كالآتي:

- ✓ قيمة معامل الارتباط تنحصر بين الصفر والواحد ، وإذا ساوى معامل الارتباط الصفر فإن المتغيرين غير مرتبطين uncorrelated ، إذا اقترب المعامل من الصفر يكون الارتباط ضعيفا ، وإذا اقترب من الواحد يكون الارتباط قويا ، وإذا ساوى معامل الارتباط الواحد يكون الارتباط تاما وهو أقوى صوره من صور الارتباط.
- ✓ إذا كانت إشارة معامل الارتباط موجبة فهذا يعني أن الارتباط طردي وإذا كانت سالبة فالارتباط عكسي .

- ✓ معامل الارتباط يقيس العلاقة غير السببية بين المتغيرين بمعنى آخر لا يهتم بتحديد من المتغير المستقل Independent ومن المتغير التابع Dependent.

ملاحظات:

- ✓ أحيانا لا يكون لمعامل الارتباط اشارة وفي هذه الحالة نهتم بدرجة العلاقة فقط.
- ✓ بعض معاملات الارتباط لها شروط لتطبيقها إذ لم تتحقق فان النتائج المتحصل عليها تكون عديمة الجدوى.
- ✓ إذا كانت المتغيرات مستقلة فان معامل الارتباط بينها يساوي الصفر والعكس غير صحيح.

(2-6) الجداول المزدوجة :Cross Tables

الجدول المزدوج هو جدول ثنائي يلخص العلاقة بين متغيرين جدوليا فيبرز الخصائص المشتركة لهما.

مثال (6-1):

الجدول التالي يمثل توزيع الطلاب والطالبات المقبولين في كلية الآداب لعام 2007 / 2008م.

جدول (6-1)

التخصص	اللغة العربية	اللغة الإنجليزية	التاريخ	الجغرافيا	الأعلام	الاجتماع	المجموع
عدد الطلاب	74	115	254	109	201	138	891
عدد الطالبات	129	134	155	57	-	485	960
المجموع	203	249	409	166	201	623	1851

يمكن تحويل التكرارات الموجودة في الجدول إلى نسب (لكل قيم الجدول) كالآتي :

- ✓ إذا تم قسمة جميع أرقام الجدول على مجموع التكرارات وهو1851نحصل على نسبة كل خلية للمجموع الكلي.
- ✓ إذا تم قسمة تكرارات أي صف على مجموع الصف سنحصل على نسبة الخلية للصف.

✓ إذا تم قسمة التكرار في كل خلية على مجموع العمود سنحصل على نسبة الخلية لمجموع العمود.

مثال (6-2):

فيما يلي قائمة ببعض البيانات التي استخرجت من ملفات بعض العاملين في إحدى المصالح الحكومية الذين يمثلون عينة عشوائية تم اختيارها من بين العاملين في تلك المصلحة الحكومية والتي تمثل مدة الخدمة للعامل والمرتب الشهري.

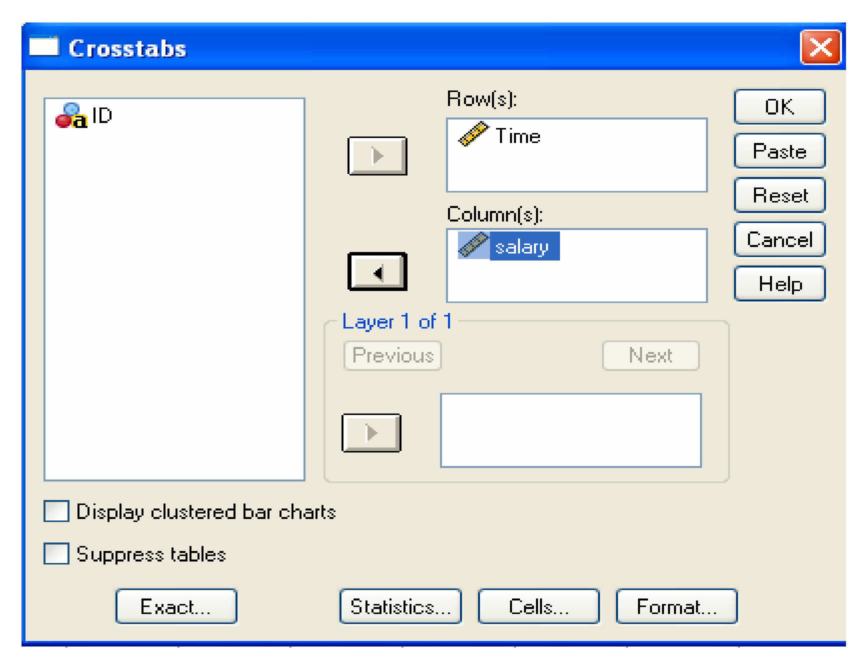
جدول (6-2)

رقم الملف	مدة الخدمة	المرتب الشهري
15/1	0.4	200
2/33	2	240
2/23	1	200
85/1	5	390
6/45	18	520
52/12	6	400
2/126	11	490
5/12	18	490
6/123	9	390
56/12	12	500
12/125	4	400
56/123	0.8	200
89/44	20	390
56/23	19	200

لإنشاء الجدول المزدوج باستخدام الحزمة SPSS لبيانات المثال نمر بالخطوات التالية:

⌖ من القائمة Analyze نختار Descriptive Statistics وننقر الأمر Cross tabs .

فيظهر الصندوق انظر الشكل(6-1) :



الشكل (6-1)

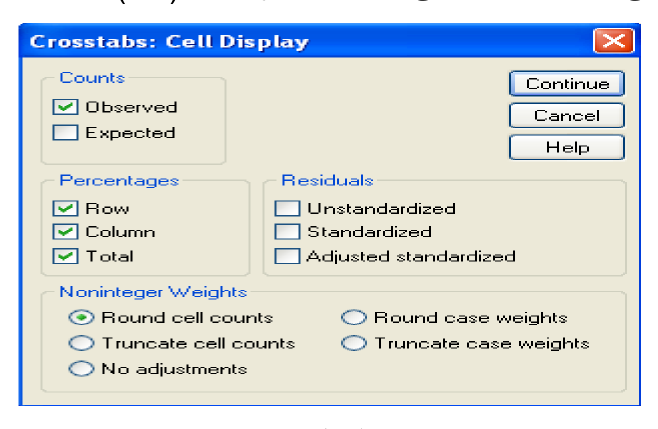
حيث أن :

- Row(s): يتضمن المتغير الذي نرغب في جعل فئاته صفوفا في الجدول.
- Column(s): يتضمن المتغير الذي نرغب في جعل فئاته أعمدة في الجدول.

Display clustered bar charts: لتمثيل بيانات الجدول بالأعمدة البيانية المزدوجة.

Suppress tables: لإخفاء الجداول المزدوجة عند التأشير.

- انقر على Cells فنحصل على الصندوق انظر الشكل (6-2):



الشكل (6-2)

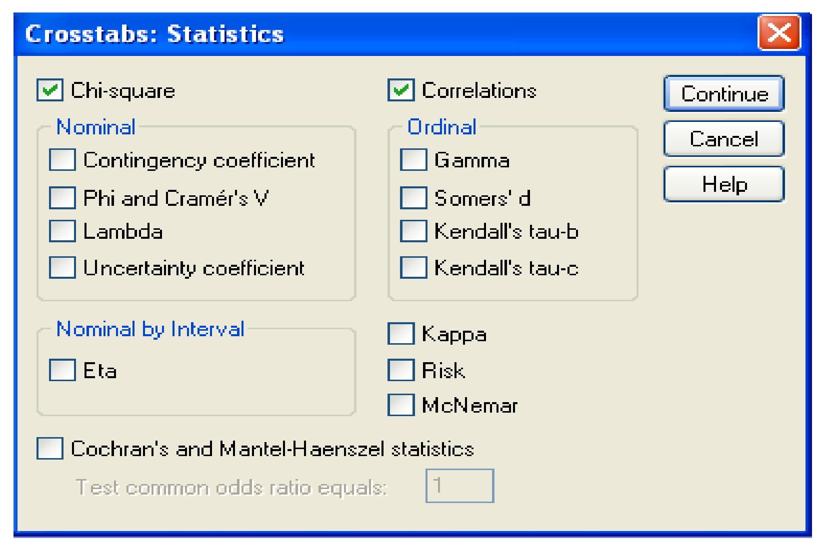
والذي نلاحظ احتوائه على العناصر التالية:

- Observed : التكرار المشاهد للخلية.
- Expected : التكرار المتوقع للخلية (يستخدم لحساب اختبار الاستقلال).
- Residuals: للحصول على البواقي.

ننقر على continue للعودة للصندوق الحواري الأصلي.

ثم ننقر على الأمر Statistics فيفتح الصندوق Crosstabls: Statistics.

انظر الشكل(3-6):



الشكل (3-6)

والذي نلاحظ احتوائه على العناصر التالية:

- Chi-square: لحساب استقلالية الصفوف عن الأعمدة.
- Correlation: يستخدم لحساب معاملي الارتباط Spearman و Pearson يحسب Spearman عندما تكون قيم المتغيرات الداخلة عددية Numeric وباعتبار أن كلا من الصفوف والأعمدة عبارة عن قيم ترتيبية.
- ويحتسب Pearson باعتبار أن كلا من الصفوف والأعمدة متغيرات كمية ويعبر عنها بالخيار Scale.

- Nominal: تتيح هذه القائمة حساب أربعة معاملات للارتباط في حالة كون كلا من الصفوف والأعمدة عوامل غير كمية بالإضافة إلى عدم امكانية ترتيب البيانات مثل متغير الجنس (ذكور/ إناث).
- Ordinal: تتيح هذه القائمة حساب أربعة معاملات للارتباط في حالة كون كلا من الصفوف والأعمدة متغيرات ترتيبية.
- Nominal by Interval: لحساب إحصائية Eta للاقتران بين متغيرين أحدهما المعتمد يقاس ضمن فترة interval scale مثل متغير الدخل والأخر مستقل له عدد محدود من الفئات مثل الجنس.

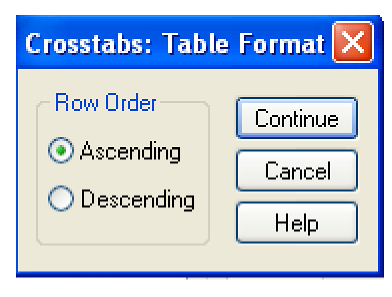
نظرا لكون الصفوف والأعمدة في هذا المثال متغيرات كمية Numeric فقد قمنا بحساب إحصائية Chi-square لاختبار الاستقلالية مع حساب قيمة معامل الارتباط بين المتغيرين.

ننقر على continue للعودة لصندوق الأصلي.

- ثم ننقر على Ok للتنفيذ.

ملاحظة:

- ✓ يمكن ترتيب صفوف الجدول تصاعديا أو تنازليا بنقر الزر Format في صندوق حوار Cross tabs كما في الشكل الآتي:



الشكل (4-6)

بعد تحديد نوعية الترتيب، هنا تم الاختيار الترتيب التصاعدي، نقوم بالنقر على Continuous لنرجع إلى القائمة الظاهرة في الشكل (6-1) وننقر على OK.

النتائج التي تم الحصول عليها في شاشة النتائج:

جدول (3-6)

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Time * salary	14	100.0%	0	.0%	14	100.0%

الجدول السابق به البيانات المتاحة والمفقودة.

جدول (3-6)

Time * salary Crosstabulation

			salary							Total
			200.00	240.00	390.00	400.00	490.00	500.00	520.00	
Time	.40	Count	1	0	0	0	0	0	0	1
		% within Time	100.0%	.0%	.0%	.0%	.0%	.0%	.0%	100.0%
		% within salary	25.0%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
		% of Total	7.1%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
	.80	Count	1	0	0	0	0	0	0	1
		% within Time	100.0%	.0%	.0%	.0%	.0%	.0%	.0%	100.0%
		% within salary	25.0%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
		% of Total	7.1%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
	1.00	Count	1	0	0	0	0	0	0	1
		% within Time	100.0%	.0%	.0%	.0%	.0%	.0%	.0%	100.0%
		% within salary	25.0%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
		% of Total	7.1%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
	2.00	Count	0	1	0	0	0	0	0	1
		% within Time	.0%	100.0%	.0%	.0%	.0%	.0%	.0%	100.0%
		% within salary	.0%	100.0%	.0%	.0%	.0%	.0%	.0%	7.1%
		% of Total	.0%	7.1%	.0%	.0%	.0%	.0%	.0%	7.1%
	4.00	Count	0	0	0	1	0	0	0	1
		% within Time	.0%	.0%	.0%	100.0%	.0%	.0%	.0%	100.0%
		% within salary	.0%	.0%	.0%	50.0%	.0%	.0%	.0%	7.1%
		% of Total	.0%	.0%	.0%	7.1%	.0%	.0%	.0%	7.1%
	5.00	Count	0	0	1	0	0	0	0	1
		% within Time	.0%	.0%	100.0%	.0%	.0%	.0%	.0%	100.0%
		% within salary	.0%	.0%	33.3%	.0%	.0%	.0%	.0%	7.1%
		% of Total	.0%	.0%	7.1%	.0%	.0%	.0%	.0%	7.1%
	6.00	Count	0	0	0	1	0	0	0	1
		% within Time	.0%	.0%	.0%	100.0%	.0%	.0%	.0%	100.0%
		% within salary	.0%	.0%	.0%	50.0%	.0%	.0%	.0%	7.1%
		% of Total	.0%	.0%	.0%	7.1%	.0%	.0%	.0%	7.1%

9.00	Count	0	0	1	0	0	0	0	1
	% within Time	.0%	.0%	100.0%	.0%	.0%	.0%	.0%	100.0%
	% within salary	.0%	.0%	33.3%	.0%	.0%	.0%	.0%	7.1%
	% of Total	.0%	.0%	7.1%	.0%	.0%	.0%	.0%	7.1%
11.00	Count	0	0	0	0	1	0	0	1
	% within Time	.0%	.0%	.0%	.0%	100.0%	.0%	.0%	100.0%
	% within salary	.0%	.0%	.0%	.0%	50.0%	.0%	.0%	7.1%
	% of Total	.0%	.0%	.0%	.0%	7.1%	.0%	.0%	7.1%
12.00	Count	0	0	0	0	0	1	0	1
	% within Time	.0%	.0%	.0%	.0%	.0%	100.0%	.0%	100.0%
	% within salary	.0%	.0%	.0%	.0%	.0%	100.0%	.0%	7.1%
	% of Total	.0%	.0%	.0%	.0%	.0%	7.1%	.0%	7.1%
18.00	Count	0	0	0	0	1	0	1	2
	% within Time	.0%	.0%	.0%	.0%	50.0%	.0%	50.0%	100.0%
	% within salary	.0%	.0%	.0%	.0%	50.0%	.0%	100.0%	14.3%
	% of Total	.0%	.0%	.0%	.0%	7.1%	.0%	7.1%	14.3%
19.00	Count	1	0	0	0	0	0	0	1
	% within Time	100.0%	.0%	.0%	.0%	.0%	.0%	.0%	100.0%
	% within salary	25.0%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
	% of Total	7.1%	.0%	.0%	.0%	.0%	.0%	.0%	7.1%
20.00	Count	0	0	1	0	0	0	0	1
	% within Time	.0%	.0%	100.0%	.0%	.0%	.0%	.0%	100.0%
	% within salary	.0%	.0%	33.3%	.0%	.0%	.0%	.0%	7.1%
	% of Total	.0%	.0%	7.1%	.0%	.0%	.0%	.0%	7.1%
Total	Count	4	1	3	2	2	1	1	14
	% within Time	28.6%	7.1%	21.4%	14.3%	14.3%	7.1%	7.1%	100.0%
	% within salary	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	28.6%	7.1%	21.4%	14.3%	14.3%	7.1%	7.1%	100.0%

الجـدول السـابق هـو الجـدول التكـراري المـزدوج للمتغيـرين الأعمـدة ل Recover والصفوف ل Treat.

- Count: هو التكرار المشاهد.
- % within treat: النسبة للصف.
- % Within recover: النسبة للعمود. ثم للكلي % of total .

جدول (4-6)

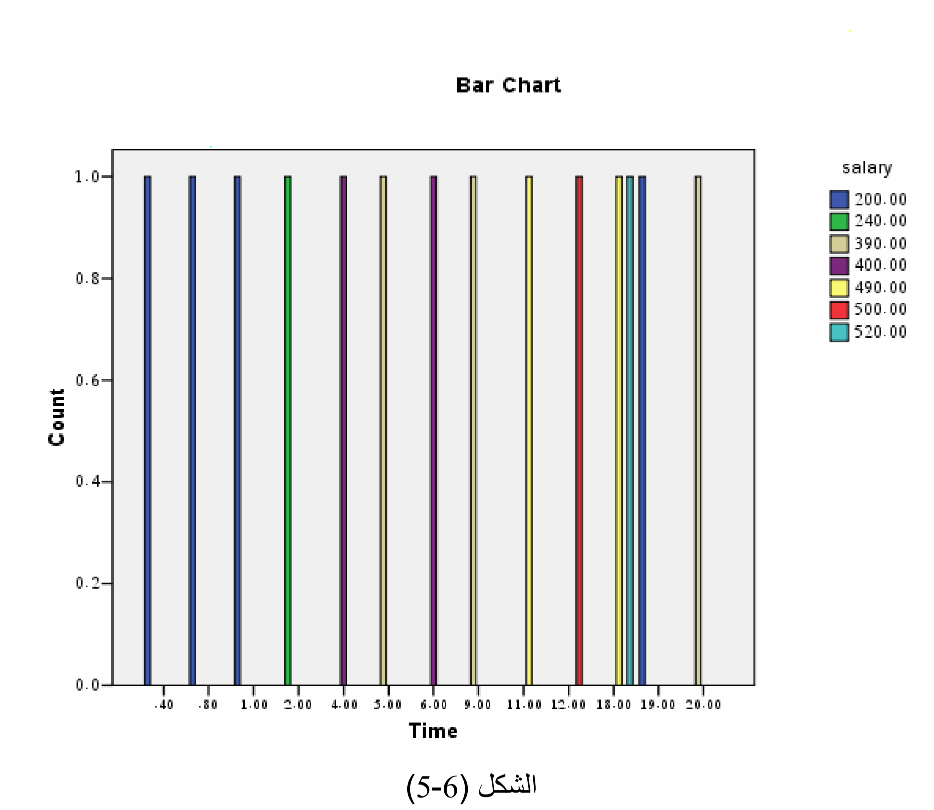
Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	73.500[a]	72	.429
Likelihood Ratio	47.894	72	.987
Linear-by-Linear Association	3.758	1	.053
N of Valid Cases	14		

a. 91 cells (100.0%) have expected count less than 5. The minimum expected count is .07.

في الجدول السابق بعد اختبار فرض العدم القائل باستقلالية الصفوف عن الأعمدة (أي استقلالية مدة الخدمة عن المرتب) ضد الفرض البديل القائل بوجود علاقة بين مدة الخدمة والمرتب ، كانت لدينا القيمة الإحصائية المحسوبة 73.5 من العلاقة :

وأن قيمة $p\text{-value} = 0.429 > 0.05$ تدعونا إلى قبول الفرض العدم بمستوى دلالة (معنوية) 0.05 أي لا توجد علاقة بين مدة الخدمة والمرتب في تلك المصلحة الحكومية.



الشكل (6-5)

- ✓ تم الحصول على رسمة الأعمدة البيانية المزدوجة للمتغيرين معا وهي أعمدة متلاصقة كل عمود مخصص لبيانات متغير معين وله لون مختلف حيث أن تكرار مدة الخدمة واحد.

(6-3) شكل الانتشار Scatter Plot:

شكل الانتشار ما هو إلا تمثيل للبيانات المزدوجة بنقاط على محورين أحدهما المحور السيني ويمثل عليه المتغير المستقل Independent والآخر المحور الصادي

ويمثل عليه المتغير التابع Dependent. وبالنظر إلى شكل الانتشار نستطيع تحديد نوع العلاقة ودرجتها تقريبا طبقا للضوابط التالية:

- إذا تجمعت النقاط حول خط يصنع زاوية حادة أو منفرجة مع المحور الأفقي تكون العلاقة خطية، وإذا تجمعت البيانات حول أي شكل آخر غير الخط تكون العلاقة غير خطية.
- إذا كانت العلاقة خطية واقتربت النقاط حول خط وهمي يمكن تخيله وتحديد ما إذا كانت العلاقة قوية والعكس صحيح .
- إذا كانت العلاقة خطية (أم غير خطية) ووقعت النقاط في شكل الانتشار جميعها فوق الخط كانت العلاقة تامة وهذه أشد صورة من صور الارتباط .

مثال 6-3:

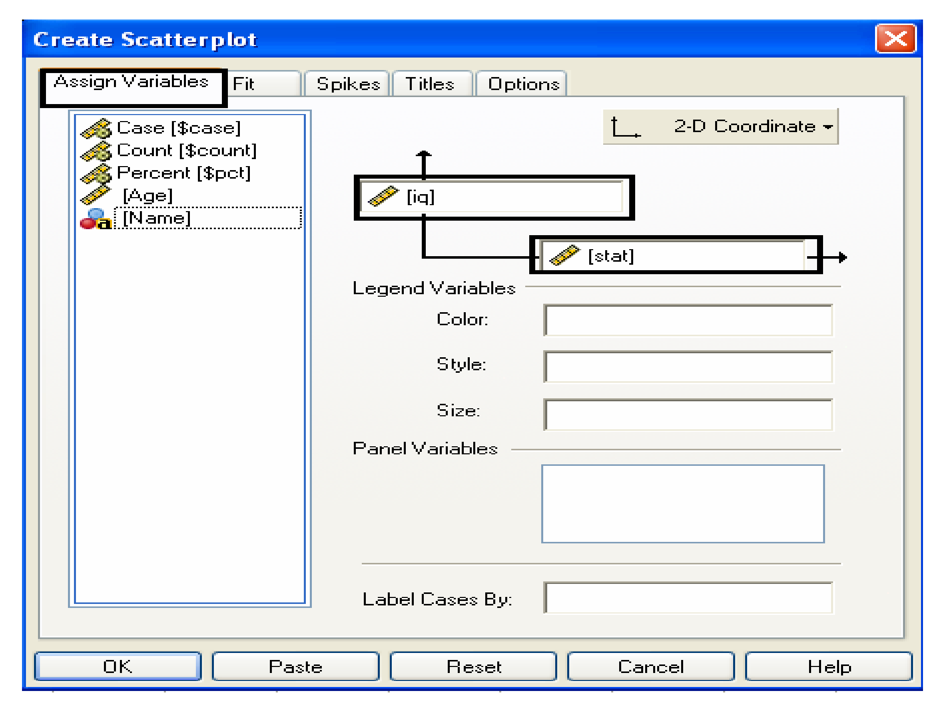
نفرض أن لدينا بيانات خاصة بنتائج 25 طالب في مقررين يرمز لها بالرموز stat، iq

جدول (6-5)

IQ	STAT	العمر	الاسم
60.00	40.00	21.00	Ahmad
70.00	59.00	30.00	Hanna
50.00	55.00	21.00	Heba
90.00	85.00	22.00	Kamala
80.00	77.00	27.00	Mervat
83.00	80.00	23.00	Mohammed
50.00	35.00	22.00	Samer
60.00	66.00	25.00	Satar
80.00	50.00	21.00	Shimaa

باستخدام البرنامج يمكن رسم شكل الانتشار كالآتي:

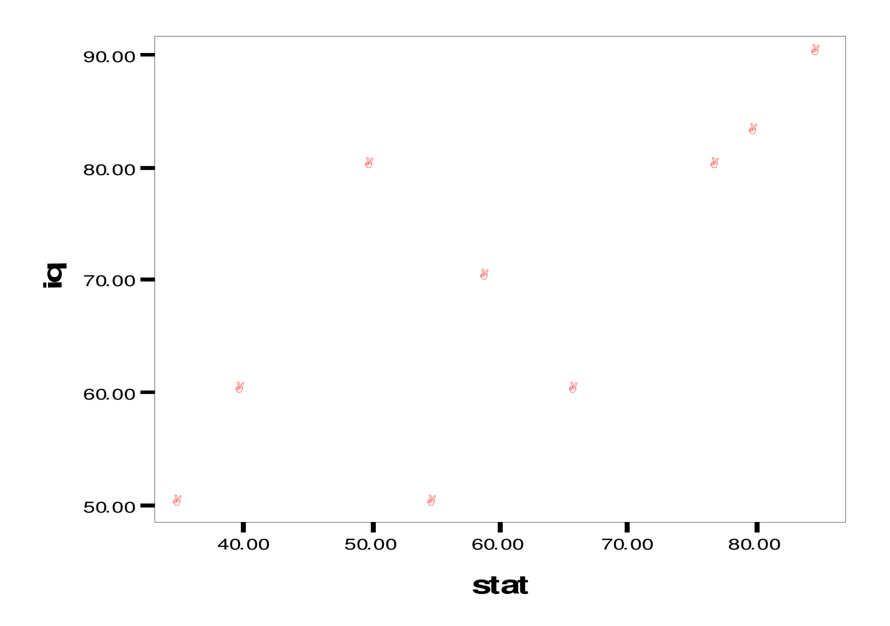
- ✓ من القائمة Graphs انقر الأمر interactive ثم انقر Scatterplot يفتح الصندوق Scatter plot انظر الشكل (6-6).



الشكل(6-6)

انقر على الأمر Assign Variables وذلك لإدخال المتغيرات التي نرغب في إيجاد شكل الانتشار لها وليكن المتغير iq المستطيل الخاص بالمحور الرأسي Y Axis، والمتغير stat الى مستطيل المحور الأفقي X Axis كما هو واضح في الشكل السابق.

- انقر على الأمر Title لكتابة عنوان الشكل المرسوم ثم انقر continue للعودة للصندوق الأصلي وليكن نريده أن يكون STAT and IQ.
- انقر على الأمر options لتحدد موقفك عن كيفية معالجة القيم المفقودة.
- انقر على الأمر continue لتعود لصندوق الحواري الأصلي.
- انقر على Ok لتنفيذ الأمر.



الشكل (6-7)

يتضح من شكل الانتشار أن البيانات متجمعة حول خط يصنع زاوية حادة مع المحور الأفقي أي أن العلاقة بين المتغيرين يمكن أن تكون خطية طردية ، وكذلك العلاقة غير تامة لأن النقاط جميعها لم تتجمع فوق خط معين ،لا يمكن معرفة قيمة معامل الارتباط من الشكل ولكن يمكن تصور اقترابه من رقم معين . وسوف نتعرض في الفصول القادمة كيفية حساب معامل الارتباط للبيانات المختلفة.

الفصــل الســـــابع

اختبارات عينـة أو عينتين

One and Two Samples Tests

(1-7) مقدمة:

تعتبر اختبارات الفروض الإحصائية واحدة من أهم التطبيقات التي قدمها علم الإحصاء كحل للمشاكل العلمية المختلفة بشتى فروع العلم .

وتبدأ أي مشكلة بحثية باهتمام الباحث بدراسة خصائص مجتمع الدراسة وتتم دراسة هذه الخصائص بعدة طرق منها اختبار فرض (أو فروض) يتعلق بها لكي يتأكد الباحث من صحة أو عدم صحة هذا الفرض حيث أنه لا يستطيع التعامل مع المجتمع ككل لان هذا إما مستحيلا أو مكلفا للغاية ولذلك ليس أمام الباحث إلا اختيار عينة عشوائية من ذلك المجتمع، ثم اختبار الفروض الموضوعة للبحث تهيئة لاستخلاص النتائج.

بصفة عامة وقبل الدخول في أية تفاصيل فإن اتخاذ أي قرار بقبول أو رفض فرضا معينا يمر بعدة مراحل وهي كالتالي:

- سحب عينة من المجتمع بحيث تكون ممثلة تمثيلاً جيداً للمجتمع.
- تجميع البيانات المتعلقة بالمشكلة من العينة مع توخي الدقة في التجميع حتى تكون النتائج قابلة للتطبيق بأقل الأخطاء الممكنة.
- تطبيق قواعد معينة لاختبار الفروض الموضوعة عن طريق الباحث وهي مشكلة تحتاج إلى خبرة إحصائية .

(2-7) تعاريف أساسية في اختبارات الفروض الإحصائية Testing Statistical Hypotheses:

فيما يلي نقدم مجموعة من المصطلحات التي تستخدم في اختبارات الفروض الإحصائية:

- **الفرض Hypothesis:**

هو تقرير أو تأكيد عن التوزيع الاحتمالي و قد يكون هذا التقرير أو التأكيد assertion متعلق بأحد معالم التوزيع الاحتمالي أو بأحد المؤشرات التي تعكس خاصية معينة من خواص التوزيع الاحتمالي.

- **الفرض العدمي H_0 (Null Hypothesis):**

فرض العدم (الفرض الصفري) نرمز له بالرمز H_0 وهو فرض حول معلمة المجتمع التي نجري الاختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة

المجتمع والإحصاءه من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما في حالة قبوله. وهذا الفرض نرفضه عندما تتوفر لدينا دلائل على عدم صحته، وتعني كلمة Null أنه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).

- **الفرض البديل H_A (Alternative Hypothesis):**

هو الفرض الذي نرمز له بالرمز H_A ويضعه الباحث كبديل عن فرض العدم و نقبله عندما نرفض الفرض العدم باعتباره ليس صحيحاً بناءً على المعلومات المستقاة من العينة.

- **اختبار الفرض الإحصائي A test of a statistical Hypothesis**

يعرف اختبار الفرض الإحصائي بأنه القاعدة التي بناء على دراسة مفردات العينة تؤدي إلى قرار معين بشأن قبول أو رفض فرض العدم. ويلاحظ أن القرار الذي نصل إليه طبقا للقاعدة المشار إليها يتعلق عادة بفرض العدم أي H_0 ، فإما أن يكون القرار هو رفض H_0 (وفي هذه الحالة يعني ذلك قبول H_A) أو قبول H_0 (وفي هذه الحالة يعني ذلك رفض H_A).

- **الخطأ من النوع الأول و الخطأ من النوع الثاني**

Type 1 error and Type 2 error

يعتمد القرار الذي يتم اتخاذه بشأن فرض العدم على المعلومات التي تعطيها العينة العشوائية التي يتم سحبها. أي أن قبول فرض العدم (رفض الفرض البديل) أو رفض فرض العدم (قبول الفرض البديل) يعتمد على النتيجة التي تعطيها العينة العشوائية، وهنا قد تعطي العينة معلومات أو نتيجة تؤدي إلى رفض فرض العدم برغم كونه في الحقيقة صحيحا، وفي هذة الحالة يكون القرار خاطئا، ومن ناحية أخرى قد تعطي العينة معلومات ِأو نتيجة في صالح فرض العدم فيتم قبوله رغم كونه في الحقيقة غير صحيح وهذا قرار خاطىء آخر.

وفي الحقيقة بسبب اعتماد القرار الذي يتم اتخاذه بشان فرض العدم على نتيجة العينة العشوائية فإن الحالات الممكنة للقرار الذي يتخذه بشأن الفرض العدم هي كما يلي:

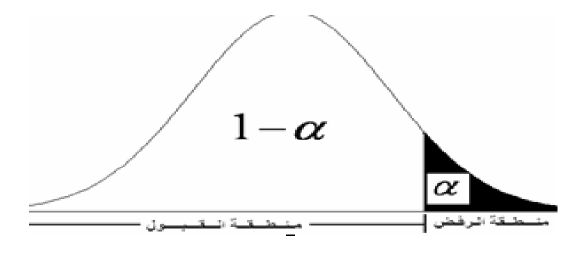
القـــرار		على مستوي المجتمع	
		H_0 صحيح	H_0 غير صحيح
على مستوى العينة	قبول H_0	القرار صحيح	خطا من النوع الثاني
	رفض H_0	خطا من النوع الأول	القرار صحيح

- **أنواع اختبارات الفروض:**

عندما نقبل فرض العدم فإننا نقبله بمستوى دقة معين فقد تكون 0.90 أو 0.95 أو 0.99 أو غير ذلك وتسمى مستويات الثقة Significance Levels أي يوجد نسبة خطأ معين في قبولنا للفرض العدم ونرمز لها بالرمز α ويسمى مستوى المعنوية، أي إذا كان مستوى الثقة 0.95 =($1-\alpha$) فان مستوى المعنوية α تساوي 0.05 وهي عبارة عن مساحة منطقة تحت منحنى التوزيع تمثل منطقة الرفض وتكون إما على صورة ذيل واحد جهة اليمين أو اليسار أو ذيلين متساويين في المساحة واحد جهة اليمين والثاني جهة اليسار.

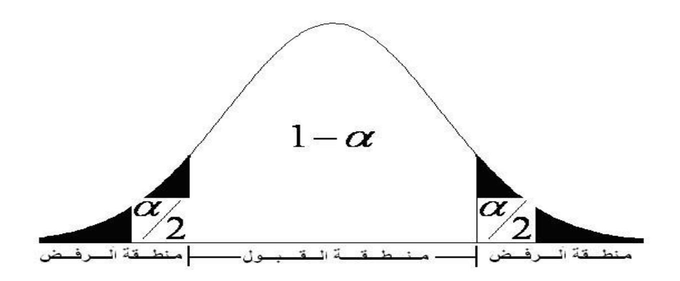
- **تعريف اختبار الفروض في جانب واحد:**

هو الاختبار الذي يضع الفرض البديل للمعلمة المجتمع علي انها أكبر أو أصغر من قيمة فرض العدم لتلك المعلمة، فتكون منطقة الرفض إما من اليمين اذا كان الفرض البديل أكبر أو من اليسار إذا كان الفرض البديل أقل.



- **تعريف اختبار الفروض في جانبين(ذيلين):**

هو الاختبار الذي يكون فيه الفرض البديل عن معلمة المجتمع على هيئة ≠ ، بل مجرد أنها تختلف. اختبار من طرفين (من الجهة اليمنى و اليسرى) $H_A : \theta \neq \theta_0$



- **خطوات اختبارات الفروض الإحصائية لحجم ثابت للعينة**

Steps for Statistical Hypotheses Testing With Fixed Sample Size

عند إجراء اختبارات الفروض الإحصائية عملياً تكون هناك خطوات محددة لإجراء الاختبار حيث تكون لازمة وبالترتيب المحدد لكل منها لإجراء الاختبار واتخاذ القرار بشأنه بطريقة سليمة. والخطوات التالية تحدد كيفية إجراء اختبارات الفروض الإحصائية التي تعتمد على حجم ثابت للعينة:

- صياغة كل من الفرض العدم H_0 والفرض البديل H_A حسب المشكلة المراد إجراء اختبارات الفروض لها.
- تحديد مستوى المعنوية للاختبار بما يلائم تكلفة الوقوع في الخطا من النوع الأول أو الخطأ من النوع الثاني.
- تحديد دالة الاختبار وهي إحصائية أي أنها تعتمد فقط على مشاهدات العينة أو الفرض العدم وتختار بحيث يكون توزيعها الاحتمالي معلومة تماما تحت فرض العدم أي أن دالة الاختبار ما هي إلا دالة في مشاهدات العينة توزيعها الاحتمالي معلوم.
- تكوين قاعدة القرار وتعتمد على كل من الفرض البديل ومستوى المعنوية ونقصد بقاعدة القرار أي هي القاعدة التي يبنى عليها اتخاذ القرار إما بقبول فرض العدم (إذا وقع خارج المنطقة الحرجة أي منطقة الرفض للاختبار) أو يتم اتخاذ قرار برفض فرض العدم (إذ وقعت داخل المنطقة الحرجة أي منطقة الرفض).

- حساب قيمة دالة الاختبار.
- اتخاذ القرار هناك طريقتين لاتخاذ قرار في الاختبارات الإحصائية وذلك إما برفض أو قبول الفرض العدمي:

الطريقة الأولى:

تعتمد على قيم محسوبة لدالة الاختبار وقيم جدولية للاختبار تستخرج من جداول أعدت خصيصا لذلك الغرض.

الطريقة الثانية:

لا تتطلب وجود جداول بل تعتمد على قيم تحسبها الحزمة SPSS أو أية حزمة إحصائية ونطلق عليها p-value ، ونستطيع تلخيص الطريقتين كالتالي :

الطريقة الأولى:

- الاختبار الإحصائي قد يكون من طرف واحد (يمين أو يسار) One Tail Test أو من طرفين Two Tail Test ، إذا كان الاختبار من طرفين نوجد قيمتين حرجتين نحصل عليهما من جداول التوزيع العيني الذي يتحكم في الاختبار أحدهما في الذيل الأيمن والأخرى في الذيل الأيسر (طرفين) للتوزيع العيني تحت الفرض العدمي هذه القيم تحجز مساحة في كل ذيل مقدارها $\alpha/2$ ،إذا وقعت القيمة المحسوبة لإحصائي الاختبار من بيانات العينة بين هاتين القيمتين نقبل الفرض العدمي إما خلاف ذلك نقبل الفرض البديل .
- إذا كان الاختبار من طرف واحد (طرف ايمن) فان منطقة الرفض تكون من جهة اليمين فنحسب القيمة الجدولية التي تحجز مساحة في الذيل الأيمن مقدارها α، إذا كانت القيمة المحسوبة اقل من القيمة الجدولية نقبل العدمي والعكس صحيح. إذا كانت منطقة الرفض جهة اليسار (طرف أيسر) نحسب القيمة الجدولية التي تحجز مساحة في الذيل الأيسر مقدارها α، إذا كانت القيمة المحسوبة اكبر من القيمة الجدولية نقبل العدمي والعكس صحيح .

الطريقة الثانية :

هناك طريقة أخرى للرفض أو للقبول استخدمت كثيرا في الآونة الأخيرة وتعتمد على احتمال محسوب يسمى*P-Value* ويرمز له في الحزمة SPSS بالرمز .Sig ، وتعرف

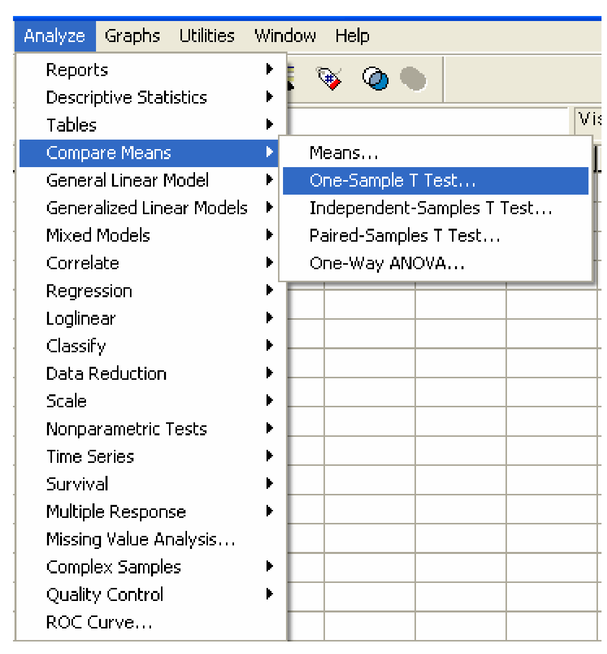
على أنها مستوى معنوية محسوب أو خطأ من النوع الأول محسوب، هذه $P\text{-}Value$ الطريقة تستخدم كثيرا في البرامج الجاهزة، تتلخص هذه الطريقة في الآتي:

- يجري الاختبار بنفس الخطوات السابق الحديث عنها ، لا تحسب قيم جدولية لإحصائي الاختبار ولكن نستخدم احتمال محسوب يرمز له بالرمز .Sig تعبر هذه القيمة عن مساحة معينة في ذيل (أو ذيلين) للتوزيع العيني لإحصائي الاختبار وتسمى قيمة مستوى المعنوية المحسوب .

إذا كان الاختبار من طرفين تقارن هذه القيمة بقيمة $\alpha/2$ ، إذا زادت القيمة المحسوبة عن($\alpha/2$) نقبل الفرض العدمي والعكس صحيح. إذا كانت الاختبار من طرف واحد نقارن القيمة المحسوبة بقيمة α ، إذا زادت عنها نقبل الفرض العدمي والعكس صحيح

.

تظهر أوامر اختبارات معلميه بالنقر على Analyze في شريط الأوامر كما في الشكل التالي:



الشكل (7-1)

(3-7) اختبار عينة واحدة

في اختبار عينة أو عينتين نحتاج إلى إحصائي للاختبار يطلق عليها اسم *t* نسبة إلى التوزيع الاحتمالي الخاص به والمسمى باسم *t-Distribution* ،التوزيع *t* من التوزيعات المتماثلة ويشبه كثيرا التوزيع الطبيعي، يؤول هذا التوزيع إلى التوزيع الطبيعي عندما يزيد حجم العينة (أكبر من 30 مفردة) معنى ذلك أنه في الحالات التي يطبق فيها إحصائي الاختبار *t* يستبدل بالتوزيع الطبيعي عندما يزيد حجم العينة عن 30 مفردة. اختبار t وهو اختبار معلمي يستخدم هذا الاختبار لفحص فرض يتعلق بالوسط الحسابي .

ويجب تحقق الشرطين التاليين حتى نستخدم اختبار t:

- يجب أن يتبع توزيع البيانات التوزيع الطبيعي، ويستعاض عن هذا الشرط بزيادة حجم العينة إلى أكثر من 30 مفردة.
- يجب أن تكون العينة عشوائية أي لا تعتمد مفرداتها على بعضها.

اختبار عينة واحدة يتوقف على فرضين أحدهما يسمى الفرض العدمي والآخر يسمى الفرض البديل، تدور هذه الفروض غالبا حول متوسط المجتمع أو تباين المجتمع أو النسبة في المجتمع ...الخ .

ولإجراء اختبار متوسط عينة نفرض أن متوسط المجتمع (المعلمة) مجهول و نريد اختبار بعض الفروض التي تدور حول المعلمة المجهولة، تحديد إحصائي الاختبار سيتوقف على هل تباين المجتمع المسحوبة منه العينة معروفا (فرض نظري لا يمكن تحقيقة) أم غير معروف. إذا كان تباين المجتمع معروفا فإن التوزيع العيني الذي يتحكم في عمليات الرفض والقبول هو التوزيع الطبيعي، إذا لم يكن تباين المجتمع معروف يكون البديل الوحيد للتوزيع الطبيعي هو توزيع t. و هو اختبار خاص بمقياس الفترات إذا كانت العينة واحدة وصغيرة أي أن حجم العينة أقل من ثلاثين، ونحاول هنا معرفة هل يختلف المتوسط الذي تم احتسابه من خلال العينة بقيمة ذات دلالة إحصائية عن متوسط المجتمع الذي أخذت منه العينة.

- الفروض الإحصائية لاختبار t:

فرض العدم : $H_0 : \mu = \hat{\mu}$

الفرض البديل: $H_A : \mu \neq \hat{\mu}$

حيث: (μ) هي متوسط المجتمع وتقرأ "ميو"

($\hat{\mu}$) هي قيمه متوسط العينه وتقرأ "ميو هات"

- **شروط استخدام توزيع t:**

- يجب أن يتبع توزيع البيانات التوزيع الطبيعي، ويعوض عن هذا الشرط بزيادة حجم العينة إلى أكثر من 30 مشاهده.

- يجب أن تكون العينة عشوائية أي لا تعتمد مشاهداتها على بعضها.

- إحصاءه t المستخدمة في الاختبار:

$$t = \frac{\overline{X} - \mu}{S / \sqrt{n}}$$

حيث: $\overline{X}$ متوسط العينه وتقرأ "اكس بار"

S الانحراف المعياري للعينة،

n حجم العينة.

- **الفرق بين اختبار t واختبار Z:**

في حالة العينات الصغيرة ($n<30$) وتباين المجتمع مجهول يتم استخدام اختبار t وفي حالة العينات الكبيرة ($n \geq 30$) يتم استخدام اختبار Z. أي أن توزيع t يؤول إلى توزيع Z عندما يزيد حجم العينة عن 30 مشاهدة.

مثال (7-1):

قررت لجنة فحص طلبية من أورق الطباعة من النوع 80 جرام التأكد من أن المتوسط الحسابي لعدد الأوراق في الرزمة الواحدة لا يساوي 500 ورقة، أخذت عينة عشوائية من الطلبية من عشرة رزم (10 رزم) وتم وزن كل زرمة فكانت النتائج كالأتي:

جدول (7-1)

الرزمة	1	2	3	4	5	6	7	8	9	10
الوزن بالكليو جرام	3.9	4.1	3.8	3.7	4	4.2	3.9	3.9	4	3.7

المطلوب ما يلي :

- هل ترفض اللجنة الطلبية بناء على هذه المعلومات عند مستوى معنوية 0.05 و0.01 .
- تقدير فترة ثقة 99% للوسط الحسابي للمجتمع .

الحل :

إذا كانت الرزمة تحتوي فعلا على 500 ورقة ووزن الورقة الواحدة 80 جرام فإن وزن الرزمة يجب أن يساوي 4 كجم (500×80 جرام = 4000 جرام) وبفرض أن وزن الرزمة يتوزع كالتوزيع الطبيعي فان قرار اللجنة يجب أن يكون مبنياً على الاختبار الإحصائي التالي:

- الفروض الإحصائية:

H_0 : متوسط وزن الرزمة = المتوسط المفترض 4.

H_A : متوسط وزن الرزمة لا يساوي المتوسط المفترض (الاختبار من طرفين).

$$H_0 : \mu = 4 \ km$$

$$\mathrm{H_A} : \mu \neq 4 \ km$$

- احصاءة الاختبار هي:

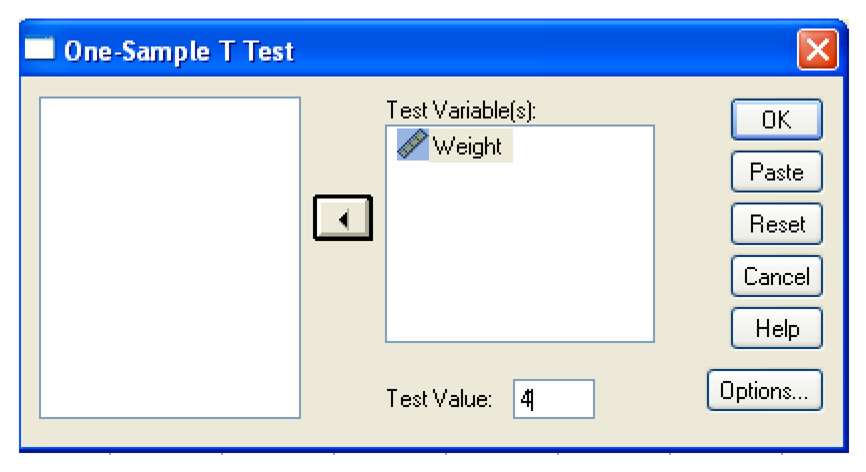
$$t = \frac{\overline{X} - \mu}{S / \sqrt{n}}$$

(لأن تباين المجتمع مجهول وحجم العينه أقل من 30)

حيث $\overline{X}$ يمثل الوسط الحسابي للعينة وS يمثل الانحراف المعياري للعينة، n حجم العينة و μ الوسط الحسابي للمجتمع بموجب فرضية العدم ويساوي 4 وهذه الإحصائية تتبع توزيع t بدرجة حرية $\upsilon = n-1$، تحتسب الإحصائية $\overline{X}$ وهي دالة الاختبار وكذالك s من قبل البرنامج) :

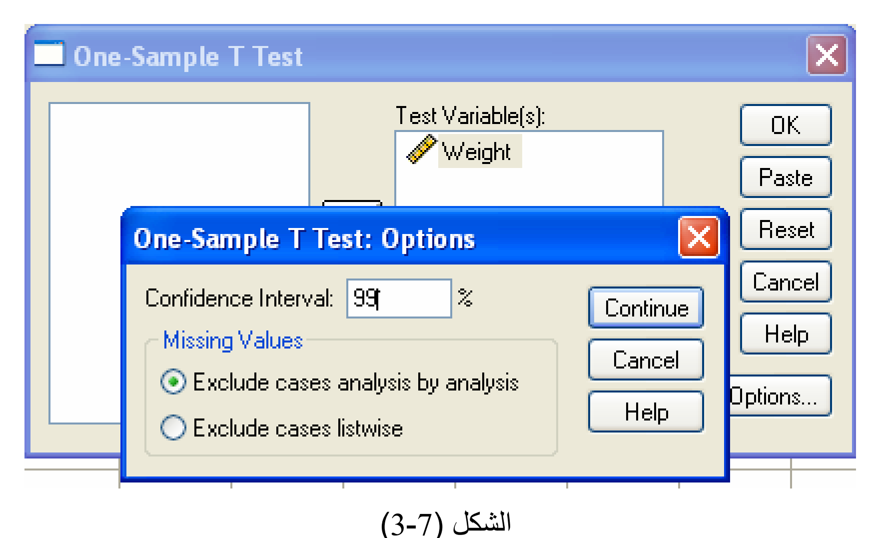
لتنفيذ الاختبار نتبع الخطوات التالية :

← من شريط القوائم نختار Analyze ومن قائمة الأوامر نختار Compare Means ونختار الأمر الفرعي One Sample t-test فيظهر صندوق حوار One Sample T-test انظر الشكل (7-2):



الشكل(7-2)

← انقر Options فيظهر صندوق حوار Options الذي نرتبه أنظر الشكل(7-3):



الشكل (7-3)

يمكن بواسطة Options تحديد فترة الثقة (المطلوب الثاني في السؤال) حيث نقوم بكتابة الفترة 99% في حقل Confidence Interval وذلك لحساب 1% مستوى معنوية فيترتب عليه درجة ثقة مقدارها المكمل لهذا الرقم أي انه 99%.

أما Missing Values فيستفاد منه في التعامل مع القيم المفقودة.

← ننقر Ok في صندوق حوار One sample T-test فتظهر النتائج التالية :

جدول (7-2)

One-Sample Statistics

Std. Error Mean	Std. Deviation	Mean	N	
.05121	.16193	3.9200	10	Weight

جدول (7-3)

One-Sample Test

	Test Value = 4					
					99% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Weight	-1.562	9	.153	-.08000	-.2464	.0864

⌖ **الطريق الأولى:**

حيث قيمه احصاءة t المحسوبة تساوي t=-1.5622 وهذه تعرف بـ t المحسوبة [$t_{calc.}$] ويجري مقارنتها بقيمة t الجدولية [$t_{tab.}$] بدرجة حرية $\nu = 9$ ولمستوى معنوية $\alpha/2$ (لأن الاختبار من طرفين) هي كالأتى

$$T = t_{(9,0.025)} = 2.262 \qquad (\alpha = 0.05)$$

$$T = t_{(9,0.005)} = 3.250 \qquad (\alpha = 0.01)$$

وحيث أن القيمة المطلقة (بدون اشارة) لـ t المحسوبة (1.5622) تقع بين حدود القيمة الجدولية ناحية اليساروااليمين ،إذن نقبل فرضية العدم ونأخذ الفرضية البديلة أي أن الوسط الحسابي للمجتمع يختلف معنويا عن القيمة 4 لمستوى معنوية 5% و1% .

⌖ **الطريقة الثانية:**

للاختبار تعتمد على قيمة p-value وتتميز عن الطريقة الأولى كونها لا تحتاج إلى استخدام جداول التوزيعات ويتم احتسابها مباشرة من قبل SPSS وهي القيمة sig.

في الجدول أعلاه ويمكن تعريف p-value بأنها أقل قيمة لـ α التي يتم عندها رفض فرضية العدم، حيث نرفض فرضية العدم عندما .sig أقل من α .

وبما أن sig.>0.05 و sig>0.01 إذن نقبل الفرض العدم لمستوى معنوية 0.05 و0.01.

أما بالنسبة لفترة ثقة 99% للفرق بين متوسط العينة ومتوسط المجتمع فيمكن كتابتها كما يلي :

$$(-0.2464 \leq \overline{X} - \mu \leq 0.0864) = 99\%$$

أما تقدير فترة ثقة 99% لمتوسط المجتمع μ كما يلي :

$$(\overline{X} - t_{(9,0.005)}\,(s/\sqrt{n}) \leq \mu \leq \overline{X} + t_{(90.005)}\,(s/\sqrt{n})) = 99\%$$

$$P(3.753578 \leq \mu \leq 4.086422) = 99\%$$

أي أن احتمال أن يقع متوسط المجتمع بين القيمتين 3.753578 و 4.086422 يساوي 99% .

وعليه نقبل الفرض العدم بمستوى معنوية 0.01 لأن القيمة 4 تقع داخل فترة الثقة وهذه تعتبر طريقة ثانية لاختبار الفرضية.

(4-7) اختبار عينتين مستقلتين Independent Grouped t-Test:

نفرض أن لدينا عينتين مستقلتين وكنا مهتمين بمتغير معين في العينتين وكنا نريد اختبار هل متوسط المجتمع المسحوب منه العينة الأولى يساوي متوسط المجتمع المسحوب منه العينة الثانية.

- الفروض الإحصائية:

فرض العدم $H_0 : \mu_1 = \mu_2$

الفرض البديل $H_1 : \mu_1 \neq \mu_2$

- عند اختيار Independent-Sample T Test يجب مراعاة الشروط التالية:
  - ✓ العينتين مستقلتين .
  - ✓ البيانات في كل عينة تتبع التوزيع الطبيعي.
  - ✓ العشوائية في اختيار العيينتين.

✓ المجتمعين المسحوبة منهما العينتين متجانس بمعنى أن تباين المجتمع المسحوب منه العينة الأولى يساوي تباين المجتمع المسحوب منه العينة الثانية ، وللتأكد من التجانس يجب إجراء اختبار سابق لاختبار T يسمى اختبار التجانس ويكون الفرض العدمي والبديل للاختبار كالتالي :

← الفرض العدمي هناك تجانس (تباين المجتمع الأول يساوي تباين المجتمع الثاني).

← الفرض البديل لا يوجد تجانس (تباين المجتمع الأول لا يساوي تبين المجتمع الثاني). والذي ممكن أن يكونا على الشكل التالي:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

أ- إذا تم قبول فرض العدم أي أن هناك تجانس سنستمر في إجراء اختبار t.

ب- إذا تم قبول الفرض البديل أي عدم وجود التجانس فإنه لا يجوز استخدام اختبار ويستبدل باختبار آخر شبيه باختبار t.

ج- اختبار التجانس يتم بواسطة اختبار آخر يسمى F-test، البرنامج يجري اختبار التجانس ويعطي نتائج اختبار t إذا كان هناك تجانس ويعطي أيضا نتائج لاختبار الشبيه باختبار t إذا لم يكن هناك تجانس وعلى المستخدم اختيار النتيجة الملائمة له، وعند تفعيل هذا الخيار يظهر الشكل التالي:

Independent-Samples T Test
Subject [subject]
Anxiety [anxiety]
Tension [tension]
Trial 1 [trial1]
Trial 2 [trial2]
Trial 3 [trial3]
Trial 4 [trial4]
Test Variable(s):
Grouping Variable:
Define Groups...
OK
Paste
Reset
Cancel
Help
Options...

الشكل (4-7)

المثال التالي يوضح كيفية تطبيق هذا الاختبار في الواقع العملي:

مثال(7-2):

أخذت عينتين عشوائيتين من مجموعة متشابهة من الأطفال وأعطي أطفال العينة الأولى غذاء رمزنا لها X و أعطي أطفال العينة الثانية غذاء آخر رمزنا لها Y وكانت الزيادة في أوزان الأطفال بالكيلو جرام في العينتين بعد مدة معينة كما في الجدول التالي:

جدول (7-4)

العينة الأولي	2.5	1.5	5.5	4.5	3.5	-
العينة الثانية	2	1.5	0.5	1.5	1	2.5

اختبر فرض عدم وجود فرق بين أثر الغذائيين " X " و " Y " في متوسط زيادة وزن الأطفال عند مستوي معنوية 0.05.

الحل :

يتضح من المعطيات أن عدد الأطفال في العينة الأولى 5 وعدد الأطفال في العينة الثانية 6 وليس من الضروري أن تتساوى أحجام المجموعتين .

لإجراء اختبار الفرق بين متوسطي عينتين نتبع الخطوات التالية:

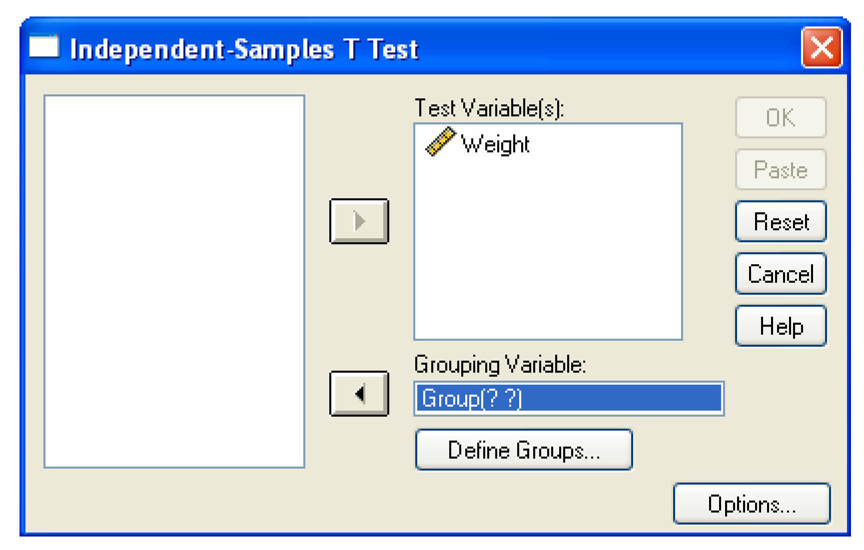
← يجب إدخال البيانات بطريقة معينة في شاشة المحرر وهي الطريقة التي يتعامل بها البرنامج كالتالي :

Weight	Group
2.50	X
1.50	X
5.50	X
4.50	X
3.50	X
2.00	Y
1.50	Y
.50	Y
1.50	Y
1.00	Y
2.50	Y

الشكل (7-5)

العمود الأول يحتوي على جميع الدرجات للطلاب و الطالبات و العمود الثاني يحتوي على A أي أن الدرجة المقابلة له من المجموعة الأولى (طلاب) و B أي أن الدرجة المقابلة له من المجموعة الثانية (طالبات) .

← من القائمة Analyze ننقر الأمر Compare Means ثم ننقر الأمر الفرعي Independent sample test ليفتح صندوق حوار Independent sample T-test الذي نقوم بترتيبه كالتالي انظر الشكل (7-6) :



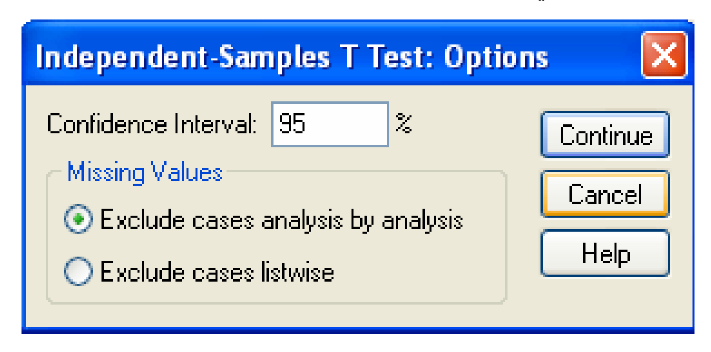
الشكل (7-6)

قمنا بنقل المتغير Weight ونقلناه إلى variable test و المتغير Group (متغير التجزئة) في قائمة Grouping variables ، ثم تعريف المجاميع X و Y عن طريق الزر Define Groups كالتالي انظر الشكل (7-7):

Define Groups
Group 1: X
Group 2: Y
Continue
Cancel
Help

الشكل (7-7)

← يمكن بواسطة النقر على Options تحديد فترة الثقة حيث نقوم بكتابة الفترة المطلوبة في حقل Confidence Interval وهنا فترة ثقة مقدرها 95% عن طريق الشكل الآتي:



الشكل (7-8)

أما Missing values فيستفاد منه في التعامل مع القيم المفقودة فتوجد طريقتين لتقدير يمكننا الاختيار بينهما.

← انقر على Continue للعودة إلى الصندوق الأصلي .

← انقر على الأمر Ok للتنفيذ .

تم الحصول على النتائج التالية :

أ- الجدول الآتي يوضح حجم المجموعة والوسط و الانحراف المعياري والخطأ المعياري لكل مجموعة .

جدول (7-5)

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Weight	X	5	3.5000	1.58114	.70711
	Y	6	1.5000	.70711	.28868

ب- الجدول التالي به اختبارين الأول اختبار التجانس والثاني اختبار T وهو كالآتي :

جدول (7-6)

Independent Samples Test

		vene's Test lity of Varia		t-test for Equality of Means						
									onfidence In the Differen	
		F	Sig.	t	df	. (2-tail	Mean fferen	td. Err fferen	Lower	Upper
Wei	Equal varia	.165	.109	.803	9	.021	0000	1362	8567	1433
	Equal varia assumed			.619	.326	.044	0000	6376	7229	2771

ويوضح الجدول من اليسار إلى اليمين الآتى :

- العمود الأول يسارا به اسم المتغير Weight و العمود الثاني والثالث يسار لإجراء اختبار التجانس كالآتي :

- فرض العدم : هناك تجانس Equal Variances.
- الفرض البديل: هناك عدم تجانس Equal Variances not assumed.
- قيمة مستوى المعنوية المحسوبة Sig.=0.109 وهي تزيد على مستوى المعنوية المحدد وهو0.05 اذن نقبل الفرض العدمي القائل بأن هناك تجانس ونرفض الفرض البديل.
- العمود الرابع والخامس لإجراء إما اختبار *t* أو الاختبار الشبية باختبار t، وحيث أننا قبلنا أن هناك تجانس نتعامل مع اختبار t والنتائج في الصف الأول من الجدول ونهمل الثاني لأنه خاص بالاختبار الشبية باختبار t .
- من الصف الأول نجد نتائج اختبار t ،قيمة مستوى المعنوية المحسوبة هي Sig.=0.021 وهي أقل من 0.05 إذن نرفض الفرض العدمي القائل أنه يوجد فرق بين مستوى الطلاب والطالبات في المجموعتين.
- يوجد أيضا حدود ثقة للفرق بين متوسطي المجتمعين يمكن للباحث استخدامها إذا رغب وكذلك خطأ الفرق .

(5-7) اختبار عينتين غير مستقلتين Paired Sample Test:

نتناول هنا موضوع اختبارات الفروض عن فرق بين متوسطي مجتمعين عندما يكون العينتين المسحوبتين منهما غير مستقلتين ويكون مشاهدات العينتين عبارة عن مشاهدات أزواج مرتبة من القيم $[(X_r, Y_r), r = 1,2,..., N]$ حيث X_1 تمثل قيمة المشاهدة للعينة الأولى مأخوذة من المفردة رقم 1 و Y_1 تمثل قيمة المشاهدة للعينة الثانية المأخوذة من المفردة رقم 1 وبصفة عامة فان X_r تمثل قيمة المشاهدة العينة الأولى المأخوذة عن المفردة رقم r بينما Y_r تمثل قيمة المشاهدة للعينة الثانية المأخوذة عن المفردة رقم r أي أنه N من المفردات وكل مفردة مسجل لها مشاهدتين (أو قيمتين أو قياسيين) أولهما للعينة الأولى و ثانيهما للعينة الثانية. مثل تسجيل قياسات ضغط الدم لكل شخص في عينة مكونة من 10 أشخاص مرة قبل تعاطي دواء معين ومرة أخرى بعد ذلك الدواء بعد أسبوعين أو ما يطلق عليه في الدراسات التجريبية الاختبار القبلي (Pre test) والاختبار البعدي (Post test) على نفس العينة. وبصفة عامة إذا كان لدينا N من أزواج المشاهدات المرتبة $(X_r, Y_r), r = 1,2,..., N$ مأخوذة على N من المفردات فإنه يمكن اعتبار أن: $X_1, X_2,..., X_N$ عينة مسحوبة من مجتمع متوسطه μ_1 وان $Y_1, Y_2,..., Y_N$ عينة مسحوبة من مجتمع متوسطه μ_2.

يستعمل هذا الاختبار لاكتشاف معنوية (دلالة) الفرق بين متوسطي متغيرين غير مستقلين لمجموعة (عينة) واحدة، معنى ذلك أن لدينا عينة واحدة وكل مفردة في العينة تعطى قراءتين القراءة الأولى تمثل العينة الأولى والقراءة الثانية تمثل العينة الثانية.

يعتمد الاختبار على إيجاد الفرق بين القراءات المتناظرة في العينتين (نرمز إلى مجموع الفرق بين العينتين بالرمز d).

ويكون صياغة الفروض كالتالي :

فرض العدم $H_0 : d = 0$

الفرض البديل $H_A : d \neq 0$

- **شروط الاختبارات للعينتين المترابطة:**

1. الشروط الأساسية لاختبار t.
2. الفرق d بين العينتين يتبع التوزيع الطبيعي وذلك إذا كان $(n < 30)$.

أما إذا كان $(n \geq 30)$ فهذا الشرط يسقط.

المثال يوضح كيفية تطبيق هذا الاختبار في الواقع العملي:

مثال(7-3) :

عقدت دورة تدريبية لمجموعة من الموظفين ولمعرفة مدى استفادة المتدربين من الدورة أجري لهم اختبارين الأول (القبلي) في بداية الدورة والثاني (البعدي) في نهايتها. وفيما يلي بيان بنتيجة الاختبارين لعينة عشوائية مكونه من تسعة من المتدربين:

جدول (7-6)

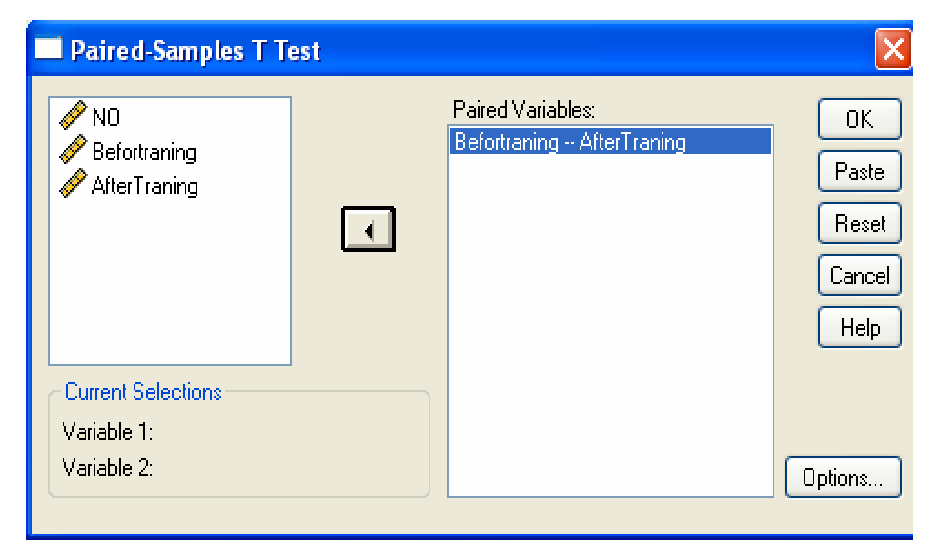
رقم المتدرب	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
الدرجة قبل التدريب	**7**	**9**	**4**	**7**	**6**	**8**	**6**	**11**	**6**
الدرجة بعد التدريب	**15**	**18**	**16**	**18**	**14**	**16**	**18**	**20**	**19**

اختبر فرض أن الدورة التدريبية ليس لها أثر في رفع مستوى المتدربين عند مستوى معنوية 5 % .

الحل :

لتنفيذ الاختبار نتبع الخطوات التالية :

← من القائمة Analyze انقر الأمر Compare Means ثم انقر الأمر الفرعي Paired – Sample T test فيظهر صندوق حوار Paired – Sample T test الذي نقوم بترتيبه انظر الشكل (7-9) :



الشكل (7-9)

- ✓ المتغيرات قبل التدريب وبعد التدريب يمثلان مشاهدات النوعين قبل التدريب وبعد التدريب حيث يتوجب إدخال كلا المتغيرين في نفس الوقت في قائمة *Paired variables*.
- ✓ يمكن بواسطة النقر على *Options* تحديد فترة الثقة حيث نقوم بكتابة الفترة المطلوبة في حقل *Confidence Interval*.
- ✓ أما *Missing values* فيستفاد منه في التعامل مع القيم المفقودة.
- ← انقر على *Continue* للعودة إلى الصندوق الأصلي .
- ← انقر على الأمر *Ok* للتنفيذ .

تم الحصول على النتائج التالية :

- ⌖ الجدول التالي يحتوي على بيانات إحصائية عن المتغيرين وتتضمن الوسط وعدد الحالات و الانحراف المعياري و الانحراف المعياري للوسط .

جدول (7-7)

Paired Samples Statistics

	Mean	N	td. Deviatio	Std. Error Mean
Pair 1 Befortran	7.1111	9	2.02759	.67586
AfterTran	17.1111	9	1.96497	.65499

⌖ الجدول التالي يحسب معامل الارتباط بين البيانات قبل وبعد ويعطي اختبار لمعامل الارتباط حيث يكون الفرض العدمي الارتباط يساوي صفر والفرض البديل الارتباط لا يساوي صفر .

معامل الارتباط موجب وهو 0.499 أما قيمة $Sig.=0.172$ وهي أكبر من 0.05 نقبل فرض العدم القائل بأن الارتباط غير معنوي.

جدول (8-7)

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 Befortraning & AfterTra	9	.499	.172

جدول (9-7)

Paired Samples Test

	Paired Differences					t	df	. (2-taile
	Mean	. Deviat	td. Erro Mean	Confidence Ir the Differenc Lower	Upper			
Pair Befortraning	0000	.00000	6667	.5373	6266	5.000	8	.000

الجدول التالي يعطي اختبار t للفرق بين عينتين غير مستقلتين ، نلاحظ أن Sig.=.000 آخر عمود وهي أقل من 0.05 إذن نرفض الفرض العدمي القائل أنه لا يوجد فرق بين العينتين ونقبل الفرض البديل القائل أن هناك فرق درجات حدث نتيجة التدريب.

هناك بيانات أخرى بالجدول يمكن للباحث الاستفادة منها إذا رغب.

الفصل الثامن

تحليل التباين

Analyzes of Variance (ANOVA)

(1-8) مقدمة:

في دراستنا لاختبارات الفروض في الفصل السابق، كانت الدراسة قاصرة على مجتمع واحد أو مجتمعين على الأكثر. وشملت تلك الدراسة اختبار فروض عن متوسط مجتمع واحد واختبارات فروض بخصوص تساوي متوسطي مجتمعين. ومن الطبيعي تعميم هذا الأسلوب إلى حالة وجود m من المجتمعات ($m \geq 2$) ويكون المطلوب منا مقارنة متوسطاتها. فمثلا إذا كان المطلوب إجراء مقارنة بين متوسطات إنتاجية الفدان لخمسة أصناف مختلفة من القطن، رمزنا لمتوسطات إنتاج الفدان من هذا الأصناف الخمسة بالرموز: $\mu_1 = \mu_2 = \ldots = \mu_5$ يكون فرض عدم وجود فرق بين هذه المتوسطات (أي فرض تساوي متوسطات هذا الأصناف الخمسة) هو:

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k \qquad (1\text{-}8)$$

ويكون الفرض البديل H_A هو أن فرض العدم ليس صحيحا أي أن H_A هو فرض عدم تساوي متوسطين على الأقل من هذه المتوسطات ويكون:

H_A: يوجد اثنين على الأقل من المتوشطات $\mu_1, \mu_2, \ldots, \mu_5$ غير متساويين (2-8)

ويستخدم أسلوب تحليل التباين في اختبار فرض تساوي متوسطات عدة مجتمعات ضد فرض عدم تساويها فمثلا يكون فرض العدم والفرض البديل كما في (8-1)، (8-2) وتتضح مدى أهمية أسلوب تحليل التباين في اختبار فرض عدم مثل H_0 في (8-1) ضد فرض بديل مثل H_A بملاحظة أنه لإجراء مثل هذا الاختبار بإتباع الأسلوب السابق عرضه (باستخدام اختبار t) لمقارنة متوسطي كل مجموعتين على حدة فإن هذا سوف يستلزم إجراء عشرة اختبارات من ذلك النوع ولاشك أن هذا سوف يحتاج إلى كثير من الوقت والجهد. ومن الواضح أنه كلما زاد عدد المجتمعات المطلوب مقارنة متوسطاتها كلما زادت صعوبة إجراء المقارنات المطلوبة بإتباع أسلوب مقارنة كل متوسطين على حدة. فمثلا إذا كان مطلوب مقارنة متوسطات عشرة مجتمعات فإن ذلك يستلزم إجراء 45 اختبار كل منها لمقارنة متوسطين فقط باستخدام اختبار t. ويستخدم أسلوب تحليل التباين لمقارنة متوسطات عدة مجتمعات في العديد من البحوث مثل مقارنة متوسطات إنتاج عدة أصناف من محصول معين، مقارنة متوسطات عدة الأيام اللازمة للشفاء من مرض معين لعدة أنواع من الأدوية ، مقارنة عدة طرق من طرق

التدريس ، مقارنة عدة طرق لإنتاج سلعة معينة ، مقارنة متوسط الدخل الشهري للأسرة بعدة محافظات وهكذا.

وفي تحليل التباين يستخدم تعبير معالجات Treatments ليعبر عن التصنيفات المختلفة المستخدمة والمطلوب مقارنة متوسطاتها فقد تكون المعالجات أنواع مختلفة من الأدوية أو قد تكون أنواع من الأسمدة أو طرق مختلفة للتدريب أو محافظات مختلفة ومطلوب إجراء مقارنات بين متوسطات ظاهرة ما باستخدام عينة عشوائية من كل منها.

وباختصار يقصد بتحليل التباين العمليات الرياضية الخاصة بتقسيم مجموع المربعات الكلي لمجموعة من البيانات إلى مصادره المختلفة وتلخص نتائج التحليل في جدول يعرف بجدول تحليل التباين ANOVA Table. إن الهدف من إجراء هذا التحليل هو اختبار فرضية تساوي متوسطات مجموعة من العينات وتعرف بالمعالجات أو المعاملات Treatment دفعة واحدة ولهذا فهو يعتبر توسيعاً لاختبار t.

- **اختبار ف:Test-F**

اختبار تحليل التباين يعتمد على إحصائي اختبار يطلق عليه اسم توزيع F نسبة إلى التوزيع الاحتمالي F Distribution.

ولاختبار مساواة متوسطات المجموعات يتم تقسيم التباين الكلي للمتغير التابع إلى مركبتين الأولى معروفة المصدر وتسمى بين المجموعات (Between Group) ومصدرها الفروق بين متوسطات المجموعات، فإذا كان هذا الجزء كبيرا فإن متوسطات المجموعات غير متساوية، والثانية داخل المجموعات (Within Group) وهي الجزء غير معروف المصدر والذي يسمى في بعض الأحيان الباقي Residuals أو الخطأ Error.

- ✓ متى نرفض الفرضية التي تقول: أن متوسطات المجموعات متساوية؟

نرفض هذه الفرضية إذا كانت نسبة التباين بين المجموعات (معروف المصدر) إلى التباين داخل المجموعات (غير معروف المصدر) كبيرا، وهذه النسبة تسمى قيمة F، فإذا كانت قيمة F كبيرة نسبيا فإن متوسطات المتغير التابع للمجموعات غير متساوية، ولكن إلى أي حد تعتبر قيمة F كبيرة حتى نرفض الفرضية التي تقول أن متوسطات المجموعات متساوية؟

✓ متى نقول أن قيمة F كبيرة نسبيا؟

إذا كانت المساحة فوقها (مستوى الدلاله Sig) أقل من المستوى المقبول لدينا α والتي غالبا تساوي 0.05 فإذا كانت قيمة Sig أقل من α فإن متوسطات المجموعات غير متساوية، وإذا كانت قيمة .Sig أكبر من $\alpha = 0.05$ فإن متوسطات المجموعات متساوية.

(2-8) تحليل التباين في اتجاه واحد (الأحادي) (One-Way ANOVA):

يقصد بتحليل التباين العمليات الرياضية الخاصة بتقسيم مجموع المربعات الكلي لمجموعة من البيانات إلى مصادره المختلفة وتلخص نتائج التحليل في جدول يعرف بجدول تحليل التباين ANOVA Table . إن الهدف من إجراء هذا التحليل هو اختبار فرضية تساوي متوسطات مجموعة من العينات وتعرف بالمعالجات أو المعاملات Treatment دفعة واحدة.

✓ متى يستخدم تحليل التباين في اتجاه واحد؟

يستخدم إذا كان لدينا عددا من العينات ونريد أن نختبر هل هذه العينات مسحوبة من مجتمعات متوسطاتها متساوية أم لا؟ بمعنى آخر هل هذه العينات مسحوبة من نفس المجتمع أم لا؟ يمكن اختبار ذلك الفرض باستخدام اختبار تحليل التباين، أي أن:

- **الفروض الإحصائية:**

الفرض الصفري

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$$

ضد الفرض البديل

H_A : يوجد متوسطان على الأقل غير متساوين

يمكن تلخيص خطوات الاختبار كالتالي:

- تحسب مجموع مشاهدات كل عينة وكذلك متوسطها ، نفرض أن $\sum_{i=1}^{N} X_{ir}$ ترمز لمجموع مشاهدات العينة r ، $\overline{X}_r = \left(\sum_{i=1}^{N} X_{ir}\right)/N$ ترمز إلى وسطها الحسابي، حيث أن $r = 1,\ldots,k$ وعدد البيانات هي N فيمكن تنظيم حساب مجموع مشاهدات كل عينة ومتوسطها باستخدام جدول البيانات الآتي:

جدول (8-1)

المشاهدات	المجتمع				
	1	2	3	...	k
1	X_{11}	X_{12}	X_{13}	...	X_{1k}
2	X_{21}	X_{22}	X_{23}	...	X_{2k}
3	X_{31}	X_{32}	X_{33}	...	X_{3k}
: :	: :	: :	: :	... ...	000 000
N_j	$X_{N_1 1}$	$X_{N_2 2}$	$X_{N_3 3}$	...	$X_{N_k k}$
$\sum$	$\sum_{i=1}^{N_1} X_{i1}$	$\sum_{i=1}^{N_2} X_{i2}$	$\sum_{i=1}^{N_3} X_{i3}$	...	$\sum_{i=1}^{N_k} X_{ik}$
$\overline{X}$	$\overline{X_1}$	$\overline{X_2}$	$\overline{X_3}$	...	$\overline{X_k}$

- يحسب المجموع الكلي لمشاهدات جميع العينات ونرمز له بالرمز $\sum$ والمتوسط العام $\overline{X}$ كآلاتي:

$$\sum = \sum_{i=1}^{N_1} X_{i1} + \sum_{i=1}^{N_2} X_{i2} + \sum_{i=1}^{N_3} X_{i3} + \ldots + \sum_{i=1}^{N_k} X_{iK}$$

وأيضا $\overline{X} = \frac{\sum}{N}$ حيث أن $N = N_1 + N_2 + \ldots + N_k$.

- يحسب مجموع المربعات الكلي Total sum of squares كآلاتي والذي نرمز له بالرمز SST:

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{N_i} X^2{}_{ji} - \frac{\sum^2}{N}$$

$$SST = X_{11}^{\ 2} + X_{21}^{2} + \ldots + X_{N_1 1}^{2} + X_{12}^{2} + X_{22}^{2} + \ldots + X_{N_2 2}^{2} + \ldots + X_{N_k k}^{2} - \frac{\sum^{\ \ 2}}{N}$$

(3-8)

أي أن مجموع المربعات الكلي نحصل عليه بإيجاد مجموع مربعات مشاهدات العينات مطروحا منه خارج قسمة مربع المجموع الكلي لمشاهدات جميع العينات على N.

ويمكن حساب مجموع المربعات الكلي بطريقة أخرى كالأتي:

$$SST = \sum_{i=1}^{k} \ \sum_{j=1}^{N_i} \left(X_{ji} - \overline{X} \right)^2$$

أي أن مجموع المربعات الكلي هو مجموع مربعات انحرافات كل مشاهدة من مشاهدات جميع العينات المأخوذة عن المتوسط العام لجميع هذه المشاهدات $\overline{X}$.

- يحسب مجموع مربعات بين المعالجات Sum of squares Between treatments

مجموع مربعات بين المعالجات والذي نرمز له بالرمز SSR :

$$SSR = \frac{\left(\sum_{i=1}^{N_1} X_{i1} \right)^2}{N_1} + \frac{\left(\sum_{i=1}^{N_2} X_{i2} \right)^2}{N_2} + \ldots + \frac{\left(\sum_{i=1}^{N_k} X_{ik} \right)^2}{N_k} - \frac{\left(\sum \ \right)^2}{N}$$

-8)

(4

ويمكن حساب مجموع مربعات بين المعالجات بطريقة أخرى كالأتي:

$$SSR = \sum_{i=1}^{K} N_i \left(\overline{X}_i - \overline{X} \right)^2$$

$$SSR = N_1 \left(\overline{X}_1 - \overline{X} \right)^2 + N_2 \left(\overline{X}_2 - \overline{X} \right)^2 + \ldots + N_k \left(\overline{X}_k - \overline{X} \right)^2$$

- يحسب مجموع مربعات الأخطاء (مربعات داخل المعالجات)

Sum of Square within treatments or sum of square of Error

يحسب بالفرق بين مجموع المربعات الكلي ومجموع المربعات بين المعالجات و الذي تم حسابهم في الخطوتين السابقتين والذي نرمز له بالرمز SSE. ويمكن حسابهم بطريقة أخرى كما يلي:

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{N_i}\left(X_{ji} - \overline{X}_i\right)^2 \qquad (5\text{-}8)$$

أي أن مجموع مربعات داخل المعالجات يعبر عن التشتت داخل العينات أي عن تباعد مشاهدات كل عينة عن وسطها الحسابي.

- يحسب متوسط مربعات بين المعالجات $\sigma^2_{Between}$ وذلك بقسمة مجموع مربعات بين المعالجات على (عدد المعالجات-1) أي علي k-1، ويسمي k-1 بدرجات حرية بين المعالجات أي أن:

$$\text{MSR}= (k\text{-}1)/\text{SSR}= \sigma^2_{Between} \qquad (6\text{-}8)$$

- يحسب متوسط مربعات داخل المعالجات σ^2_{within} وذلك بقسمة مجموع مربعات داخل المعالجات على (عدد المشاهدات الكلية-عدد المعالجات) أي على N-K، ويسمي N-K بدرجات حرية داخل المعالجات أي أن:

$$\text{MSE} = (\text{ N-K})/\text{SSE}= \sigma^2_{within} \qquad (7\text{-}8)$$

- يحسب خارج قسمة متوسط مربعات بين المعالجات على متوسط مربعات داخل المعالجات و إذا رمزنا بذلك بالرمز F للنسبة الناتجة فإن:

$$F = \frac{\sigma^2_{Betwen}}{\sigma^2_{within}} \qquad (8\text{-}8)$$

ويكون لها توزيع F بدرجات حرية k-1 ، N–k. وتعتبر F المحسوبة هي إحصاءة الاختبار.

- إذا كانت F المحسوبة أكبر من $F_{(\alpha,k-1,n-k)}$ يرفض الفرض العدم H_0 و فيما عدا ذلك يقبل فرض العدم H_0.

والجدول الآتي الذي يعرض ملخص الخطوات السابقة يسمي جدول تحليل التباين analysis of variance table ويكون له الصيغة العامة المبينة في الجدول التالي:

جدول (8-2)

مصدر الاختلاف	مجموع المربعات SS	درجات الحرية d.f	متوسط المربعات MS	F
بين المعاملات	SSR	K-1	MSR	$F_{calc} = \frac{MSR}{MSE}$
الخطأ	SSE	$\sum_{i=1}^{K} N_i - k$	MSE	
المجموع	SST	$\sum_{i=1}^{K} N_i - 1$		

والآن نتناول كيف يمكننا تنفيذ هذا باستخدام البرنامج مما يعمل علي إتاحة استخدام ذلك مع توفير الوقت والجهد في حالة زيادة عدد بيانات التحليل الإحصائي وذلك من خلال أمثلة تطبيقية متنوعة:

مثال(8-1):

استخدمت ثلاث طرق تعليمية مختلفة في تدريس مادة الحساب لثلاث عينات متشابهة من الطلاب وكانت درجاتهم في الامتحان النهائي كما بالجدول التالي:

جدول (8-3)

العينة الأولى	**8**	**7.5**	**7**	**5.5**
العينة الثانية	**9**	**10**	**8**	**9**
العينة الثالثة	**6**	**5**	**5.5**	**3.5**

اختبر ما إذا كان هناك فروق معنوية بين متوسطات درجات الطلبة في الطرق الثلاث عند مستوى معنوية 0.05.

الحل:

في البداية يتم إدخال البيانات كلها في البرنامج في عمود واحد مع إضافة عمود ثاني يشير إلى رقم العينة فيكون شاشة data editor وقد سمناه factor نضع فيه قيمة

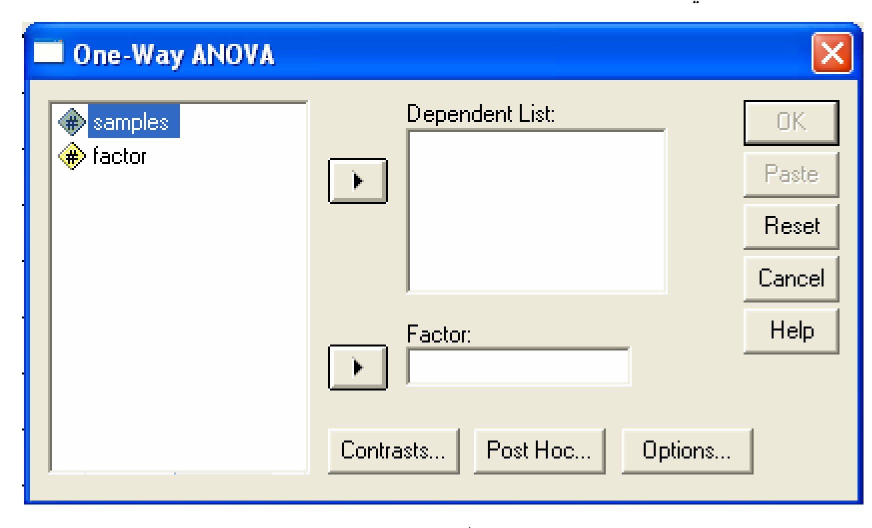
1 لرمز أن القيمة الموجودة في العمود samples هي قيمة العينة الأولـى ونضـع القيمـة 2 لرمز أن القيمة الموجودة في العمود samples هي قيمة العينة الثانية وهكذا ونضـع القيمة 3 لرمز أن القيمة الموجودة في العمود samples هـي قيمـة العينـة الثالثـة وتكـون البيانات كما في الشكل التالي:

samples	factor
8.00	1
7.50	1
7.00	1
5.50	1
9.00	2
10.00	2
8.00	2
9.00	2
6.00	3
5.00	3
5.50	3
3.50	3

وبعد إدخال البيانات نتبع الخطوات التالية:

- اختيار Compare means.
- اختيار One-Way ANOVA.

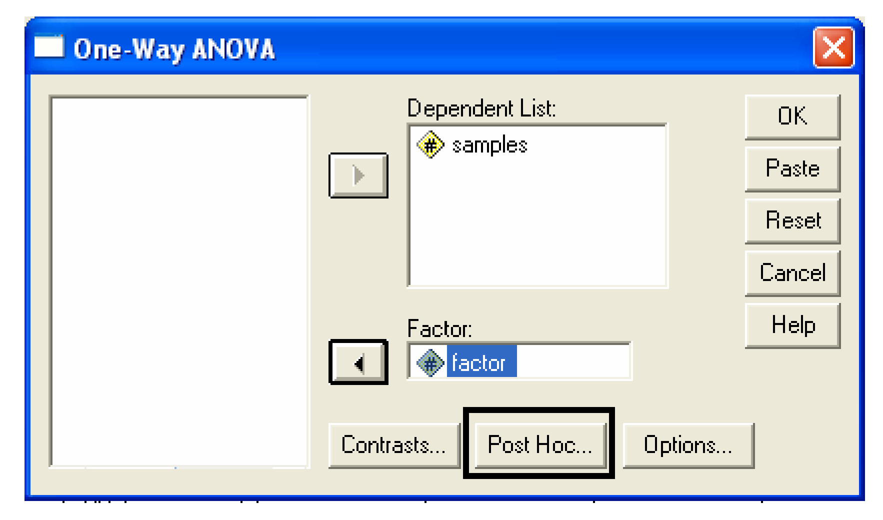
فيظهر الشكل الاتي:



شكل (8-1)

⌖ ننقل المتغيرات المراد حساب الفروق فيها إلى Dependent List ، ومتغير تصنيف المشاهدات في Factor.

فيظهر الشكل الآتي:



شكل (8-2)

⌖ ننقر على OK تظهر النتائج كما في الجدول التالي:

جدول (8-4)

ANOVA

SAMPLES

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	32.000	2	16.000	16.000	.001
Within Groups	9.000	9	1.000		
Total	41.000	11			

القرار الإحصائي:

يتضح من الجدول أن F = 16.00 عند 0.05 وهي أقل من P-value = 0.001 وهذا يدل على وجود فروق بين الطرق الثلاث لتدريس المهارات الحسابية.

نظراً لوجود فروق معنوية بين متوسطات الطرق فهذا يعني عدم تساوي متوسطي طريقتين على الأقل ولاختبار معنوية الفروق لكل زوج من المعالجات نلجأ إلى المقارنات المتعددة باستخدام طريقتي L.S.D و بإتباع الخطوات التالية:

باستخدام الخطوات السابقة حتى الشكل رقم (8-2) فنقوم بنقر على post Hoc المحددة في الشكل (8-2) فيظهر الشكل (8-3):

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

LSD, Bonferroni, Sidak, Scheffe, R-E-G-W F, R-E-G-W Q, S-N-K, Tukey, Tukey's-b, Duncan, Hochberg's GT2, Gabriel, Waller-Duncan, Type I/Type II Error Ratio: 100, Dunnett, Control Category: Last, Test: 2-sided, < Control, > Control

Equal Variances Not Assumed

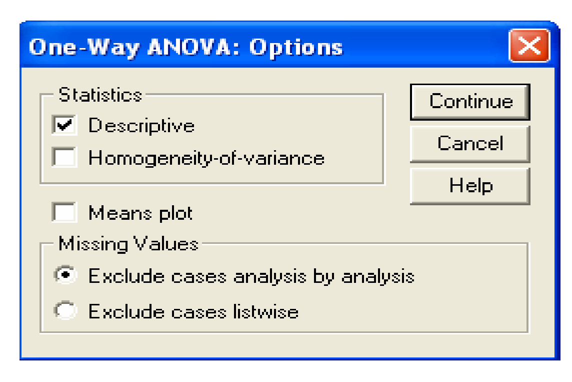
Tamhane's T2, Dunnett's T3, Games-Howell, Dunnett's C

Significance level: .05

Continue, Cancel, Help

شكل (8-3)

و نلاحظ وجود طرق عديدة للمقارنات المتعددة عندما يكون التباين لكل عينة متساوي حيث قمنا بتأشير LSD، ويوجد طرق للمقارنة عندما يكون التباين غير متساوي، وأخيراً تحديد مستوى المعنوية المطلوب 0.05 في Significance Level، عند نقر زر Continue كما نلاحظ أنه يوجد زر في الشاشة الرئيسة شاشة (8-2) يطلق عليها Options و عند تفعيل هذا الاختيار نتمكن من حساب الإحصائيات الوصفية للقيم الفعلية أو القيم المفقودة كما في الشكل التالي:



شكل (8-4)

ثم زر Ok نحصل على المخرجات التالية:

جدول (8-5)

Multiple Comparisons

Dependent Variable: SAMPLES
LSD

(I) FACTOF	(J) FACTOI	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interv Lower Bound	Upper Bound
1	2	-2.0000*	.707	.020	-3.5996	-.4004
	3	2.0000*	.707	.020	.4004	3.5996
2	1	2.0000*	.707	.020	.4004	3.5996
	3	4.0000*	.707	.000	2.4004	5.5996
3	1	-2.0000*	.707	.020	-3.5996	-.4004
	2	-4.0000*	.707	.000	-5.5996	-2.4004

*. The mean difference is significant at the .05 level.

Post Hoc Tests

نلاحظ أنه باستعمال طريقة LSD فقد ظهرت فروق معنوية بمستوى معنوية 0.05 بين متوسطات المعالجات (1و2) –(1و3) و (2و3) حيث كانت قيمة p-value أو sig. أقل من 0.05 مع وجود علامة (*) على فروق المتوسطات ووجود تلك العلامة علامة وجود فروق بين العينات.

مثال (8-2):

لمقارنة أربعة أنواع من الأدوية d، c، b،a لعلاج مرض معين اختيرت أربع عينات عشوائية كل منها مكون من خمسة أشخاص من المصابين بذلك المرض وأعطيت كل عينة نوع من الأدوية وكان عدد أيام العلاج اللازمة حتى الشفاء للأشخاص في العينات الأربعة كما بالجدول التالي:

جدول (8-6)

العينة الأولى (دواء a):	7	3	4	6	5
العينة الثانية (دواء b):	6	5	9	8	7
العينة الثالثة (دواء c):	3	2	5	6	4
العينة الرابعة (دواء d):	8	6	3	7	6

اختبر فرض عدم وجود فرق معنوي بين متوسطات عدد الأيام اللازمة حتى الشفاء للأدوية الأربعة a، b، c، d وذلك عند مستوى معنوية 0.01.

الحل:

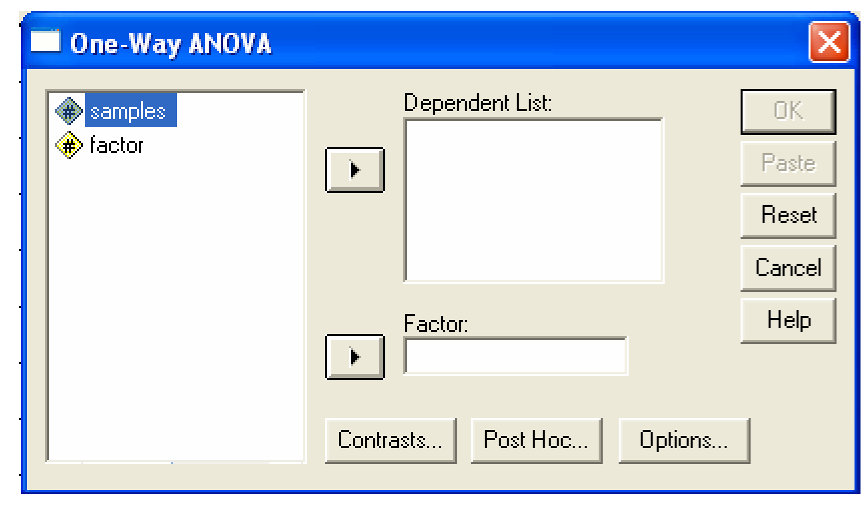
في البداية يتم إدخال البيانات كلها في البرنامج في عمود واحد مع إضافة عمود ثاني يشير إلى رقم العينة فيكون شاشة data editor وقد سمناه factor نضع فيه قيمة 1 لرمز أن القيمة الموجودة في العمود samples هي قيمة العينة قيمة الدواء الأولى a ونضع القيمة 2 لرمز أن القيمة الموجودة في العمود samples هي قيمة الدواء الثاني b وهكذا ونضع القيمة 3 لرمز أن القيمة الموجودة في العمود samples هي قيمة الدواء الثالث c وهكذا ونضع القيمة 4 لرمز أن القيمة الموجودة في العمود samples هي قيمة الدواء الرابع d وتكون البيانات كما في الشكل التالي:

samples	factor
7.00	1
3.00	1
4.00	1
6.00	1
5.00	1
6.00	2
5.00	2
9.00	2
8.00	2
7.00	2
3.00	3
2.00	3
5.00	3
6.00	3
4.00	3
8.00	4
6.00	4
3.00	4
7.00	4
6.00	4

وبعد إدخال البيانات نتبع الخطوات التالية:

- اختيار Compare means.
- اختيار One-Way ANOVA.

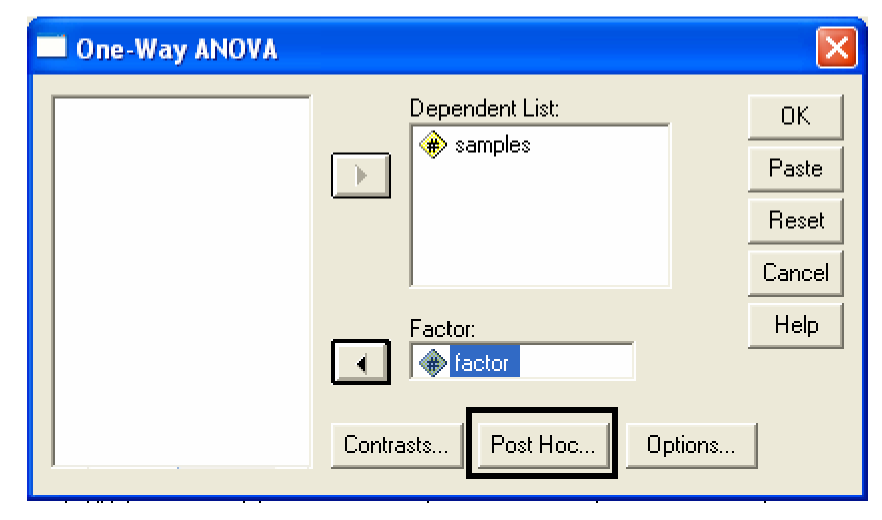
فيظهر الشكل الاتي:



شكل (5-8)

⌖ ننقل المتغيرات المراد حساب الفروق فيها إلى Dependent List، ومتغير تصنيف المشاهدات في Factor.

فيظهر الشكل الآتي:



شكل (6-8)

⌖ ننقر على OK تظهر النتائج كما في الجدول التالي:

جدول (7-8)

ANOVA

SAMPLES

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	25.000	3	8.333	3.030	.060
Within Groups	44.000	16	2.750		
Total	69.000	19			

القرار الإحصائي:

يتضح من الجدول أن p-value = 0.06 وهي أكبر من 0.01 عند F = 3.03 وهذا يدل على عدم وجود فروق بين الأنواع الأربع من الأدوية ويؤكد ذلك جدول تحليل الفروق التالية:

Post Hoc Tests

جدول (8-8)

Multiple Comparisons

Dependent Variable: SAMPLES
LSD

(I) FACTOI	(J) FACTO	Mean Difference (I-J)	Std. Error	Sig.	9% Confidence Inter	
					Lower Bound	Upper Bound
1	2	-2.0000	1.049	.075	-5.0633	1.0633
	3	1.0000	1.049	.355	-2.0633	4.0633
	4	-1.0000	1.049	.355	-4.0633	2.0633
2	1	2.0000	1.049	.075	-1.0633	5.0633
	3	3.0000	1.049	.011	6.334E-02	6.0633
	4	1.0000	1.049	.355	-2.0633	4.0633
3	1	-1.0000	1.049	.355	-4.0633	2.0633
	2	-3.0000	1.049	.011	-6.0633	6.334E-02
	4	-2.0000	1.049	.075	-5.0633	1.0633
4	1	1.0000	1.049	.355	-2.0633	4.0633
	2	-1.0000	1.049	.355	-4.0633	2.0633
	3	2.0000	1.049	.075	-1.0633	5.0633

نلاحظ أنه باستعمال طريقة LSD لم تظهر فروق معنوية بمستوى معنوية (دلالة) 0.01 بين متوسطات المعالجات كافة حيث كانت قيمة P-Value أو .Sig أكبر من 0.05 مع عدم وجود علامة (*) عند فروق المتوسطات .

(3-8) تحليل التباين في اتجاهين (الثنائي) (Two Way ANOVA):

إن تحليل التباين في اتجاه واحد يستخدم لدراسة أثر عامل واحد فقط على متغير ما ولكن عندما نريد دراسة أثر عاملين أو أكثر في هذه الحالة يمكننا استخدام تحليل التباين في اتجاهين (أو ما يسمى تحليل التباين الثنائي).

فتحليل التباين الثنائي Two Way ANOVA يمكن استخدامه لدراسة أثر متغيرين عاملين نسمي أحدهما الصفوف (المعالجات)، والآخر الأعمدة (القطاعات) ويمكن وصف البيانات تحت هذه التجارب كالتالي:

جدول (8-9)

الصفوف (Rows)		الأعمدة (Columns) القطـــــاعـــات					
		1	2		j	…	c
	1	X_{11}	X_{12}	...	X_{1j}	...	X_{1c}
	2	X_{21}	X_{22}	...	X_{2j}	...	X_{2c}
	.	.	.	...	.	...	.
	.	.	.	...	.	...	.
	.						
معالجات	i	X_{i1}	X_{i2}	...	X_{ij}	...	X_{ic}
	.	.	.	...	.	...	.
	.						
	.	.	.	...	.	...	.
	.	.	.	...	.	...	.
	.	.	.	…	.	…	.
	r	X_{r1}	X_{r2}	...	X_{rj}	....	X_{rc}

⌖ يمكن صياغة الفروض الإحصائية في هذه الحالة كالتالي:

(أ) للمعالجات:

$$H_0^{(1)} : \mu_1 = \mu_2 = \ldots = \mu_r$$

ضد الفرض البديل

$H_A^{(1)}$: يوجد متوسطان على الأقل غير متساوين

(ب) للقطاعات:

$$H_0^{(2)} : \mu_1 = \mu_2 = \ldots = \mu_c$$

ضد الفرض البديل

يوجد متوسطان على الأقل غير متساوين : $H_A^{(2)}$

مثال (8-3)

إذا كان لدينا أربعة أنواع من الأسمدة A،B،C،D استخدمت كمعالجات لأربع قطاعات

القطاع المعالج	القطاع الأول	القطاع الثاني	القطاع الثالث	القطاع الرابع	المجموع
A	9.3	9.4	9.6	10	38.3
B	9.4	9.3	9.8	9.9	38.4
C	9.2	9.4	9.5	9.7	37.8
D	9.7	9.6	10	10.2	39.5
المجم	37.6	37.7	38.9	39.8	154

مختلفة، حيث كانت النتائج كما في الجدول التالي:

استخدم تحليل التباين في اتجاهين لمعرفة ما إذا كان لها تأثير مختلف على زيادة إنتاج القمح أم لا؟

الحل:

أولا الفروض الإحصائية بالنسبة للمعالجات:

H_0 : لا يوجد فروق بين متوسطات المعالجات (الصفوف).

H_1 : يوجد فروق بين متوسطات المعالجات على الأقل لاثنين منهم.

ثانيا الفروض الإحصائية بالنسبة للقطاعات:

H_0 : لا يوجد فروق بين متوسطات القطاعات (الأعمدة).

H_1 : يوجد فروق بين متوسطات القطاعات على الأقل لاثنين منهم.

ندخل البيانات كما في الشكل التالي:

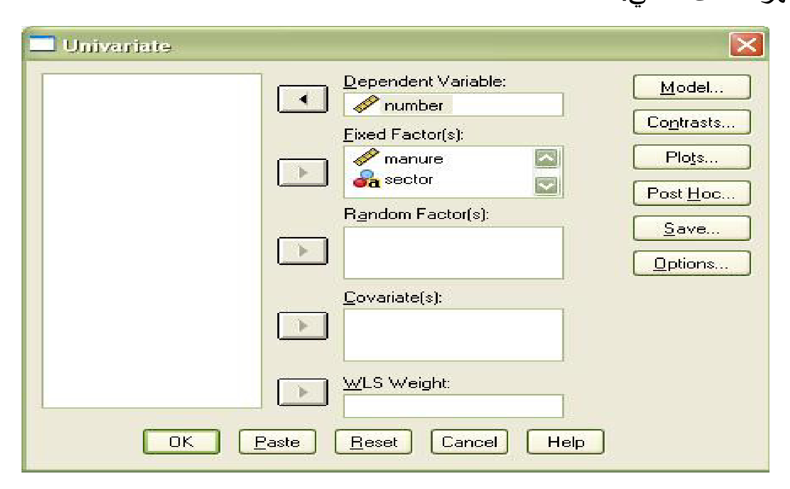
	manure	sector	number
1	1.00	A	9.30
2	1.00	B	9.40
3	1.00	C	9.20
4	1.00	D	9.70
5	2.00	A	9.40
6	2.00	B	9.30
7	2.00	C	9.40
8	2.00	D	9.60
9	3.00	A	9.60
10	3.00	B	9.80
11	3.00	C	9.50
12	3.00	D	10.00
13	4.00	A	10.00
14	4.00	B	9.90
15	4.00	C	9.70
16	4.00	D	10.20

الشكل (8-7)

ثم نتبع الآتي:

Analyze → General linear Model → Univariate

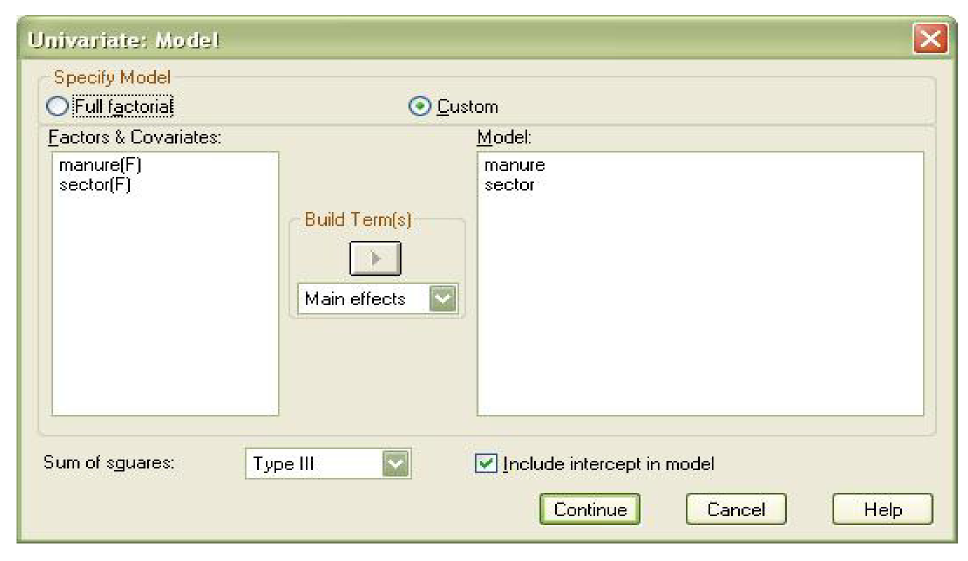
ليظهر الشكل التالي:



الشكل (8-8)

- أدخلنا العاملين في Fixed Factors على اعتبار أن كافة معالجات العامل قد ضمنت في التجربة أما في حالة أخذ عينة من معالجات العامل فإننا سنستعمل خانة Random Factor.
- انقر زر Model فيظهر صندوق حوار Model، اختر Custom بدلا Full Factorial وذلك لأننا لا نرغب في ظهور التفاعل Interaction في جدول تحليل التباين (Manure*Sector) لعدم وجود درجات حرية كافية

للخطأ التجريبي حيث يظهر صندوق حوار Model بعد ترتيبه كما في الشكل (8-9):



الشكل (8-9)

حيث أن تأشير Include Intercept In Model يعمل على تضمين الحد الثابت في النموذج الخطي العام باعتباره نموذج انحدار.

أما خانة Build Terms فتستعمل لتعيين نوع التأثيرات Effects التي يراد إظهارها في جدول تحليل التباين حيث قمنا بنقل المتغيرين manure و sector من خانة Factors & Covariates إلى خانة Model، لتظهر كتأثيرات رئيسية Main Effect (بعد التأكد من أن خانة Build Effects يتضمن الخيار Main effect ثم انقر على Continue ثم Ok في صندوق حوار Univariate.

ثم تحصل على المخرجات التالية كما في الشكل التالي:

جدول (8-10)

Between-Subjects Factors

		N
manure	1.00	4
	2.00	4
	3.00	4
	4.00	4
sector	A	4
	B	4
	C	4
	D	4

Tests of Between-Subjects Effects

Dependent Variable: number

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1.210[a]	6	.202	22.687	.000
Intercept	1482.250	1	1482.250	166753.1	.000
manure	.825	3	.275	30.937	.000
sector	.385	3	.128	14.437	.001
Error	.080	9	.009		
Total	1483.540	16			
Corrected Total	1.290	15			

a. R Squared = .938 (Adjusted R Squared = .897)

الجدول الأول يعطي بيان بعدد الحالات و المفقود، الجدول الثاني يبين مجموع المربعات بين المعالجات = 0.385 و مجموع المربعات بين القطاعات = 0.825 مجموع مربعات الخطأ = 0.08 و مجموع المربعات الكلي = 1.29.

- **القرار الإحصائي للمعالجات:**

و أيضا يتضح معنوية الفرق بين الأسمدة وكذلك بين القطاعات بمستوى معنوية 0.05 حيث أن

p-value=0.0 عند صف الأسمدة manure وهي أقل من 0.05 أي نرفض H_0 ونقبل H_A ونقول أنه يوجد فروق بين متوسطات المعالجات .

- **القرار الإحصائي للقطاعات:**

وأيضا p-value=0.001 عند صف القطاعات sector وهي أقل من 0.05 نرفض H_0 ونقبل H_A ونقول أنه يوجد فروق بين متوسطات القطاعات.

الفصل التاسع

تحليل الارتباط

Correlation Analysis

(9-1) مقدمة:

تعرضنا في الفصول السابقة لدراسة ظاهرة واحدة كالاستهلاك، الدخل، المبيعات، العمر، الوزن وبدأنا ببيان كيفية عرض البيانات التي جمعناها في جدول أو أشكال بيانية مختلفة واستطردنا بعد ذلك إلى كيفية تلخيص تلك البيانات من خلال مقاييس النزعة والتشتت لمجموعة واحدة أو مجموعات مختلفة وكيفية تقدير تلك المقادير وكيفية اختبارات الفروض عليها و الآن نبحث كيفية إيجاد العلاقات الرياضية التي تربط المتغيرات ببعضها وما مقدار هذا الارتباط من خلال ما يسمي بمقاييس الارتباط المختلفة والتي سوف نتعرض لها في هذا الفصل كما نناقش في الفصل التالي كيف يمكن التنبؤ بأحد المتغيرات لقيمة محددة للمتغير الأخر و هي ما تسمي معادلة الانحدار.

نستطيع القول بأن تحليل الارتباط والانحدار هو أداة إحصائية نستفيد منها في تحديد العلاقة بين متغيرين أو أكثر للتنبؤ بأحد المتغيرات استنادا إلى قيم المتغير أو المتغيرات الأخرى. فمثلا إذا علمنا العلاقة بين مصروفات الدعاية و المبيعات، فيمكننا الاستفادة من تحليل الارتباط للتنبؤ بالمبيعات حالما تتوفر لنا قيمة نفقات الدعاية.

وسوف نتناول في هذا الفصل قضية الارتباط بين متغيرين سواء كانوا كميين أو غير ذلك بينما في الفصل القادم نتناول كيفية التنبؤ بأحد المتغيرات في ضوء متغير آخر أو أكثر من خلال دارسة الانحدار

(9-2) الارتباط الخطى البسيط Simple Correlation

تسمي العلاقة بين ظاهرتين بالارتباط Correlation مثلاً العلاقة بين الدخل والاستهلاك فمن البديهي أن زيادة دخل الفرد يؤدي إلى زيادة استهلاكه من السلع والخدمات (علاقة طردية) كما أن ارتفاع سعر سلعة ما يؤدي إلى تدني الطلب عليها (علاقة عكسية) علماً أن الارتباط قد يكون خطياً Linear أو غير خطي Non Linear. إن المقياس المستخدم الذي يقيس درجة الارتباط يعرف بمعامل الارتباط Correlation Coefficients ويرمز له r وتتراوح قيمته بين -1 إلى 1 $[-1 \le r \le 1]$.

ومن الأمثلة على ذلك:

- الإنفاق، والدخل العائلي.
- سعر السلعة، والكمية المطلوبة منها.

- الفترة الزمنية لتخزين الخبز، وعمق طراوة الخبز.
- تقديرات الطلاب في مقرر الإحصاء، وتقديراتهم في مقرر الرياضيات.
- كميات السماد المستخدمة، وكمية الإنتاج من محصول معين تم تسميده بهذا النوع من السماد.
- عدد مرات ممارسة نوع معين من الرياضة البدنية، ومستوى الكلسترول في الدم.
- وزن الجسم، وضغط الدم.

يحسب معامل الارتباط الخطي البسيط بافتراض وجود علاقة خطية بين اثنين من المتغيرات فقط مع العلم أن الحصول على قيمة صغيرة (قريبة من الصفر) لهذا المعامل لا يعني عدم وجود علاقة بين المتغيرين فقد توجد علاقة من الدرجة الثانية (ارتباط غير خطي). ويختلف نوع المقاييس الذي نستخدمه في حساب معامل الارتباط طبقا لنوع البيانات وسوف يجرى حسابه في حالة البيانات الكمية ، والبيانات الوصفية المقاسة بمعيار ترتيبي.

- **الغرض من تحليل الارتباط الخطى البسيط**

الغرض من تحليل الارتباط الخطي البسيط هو تحديد نوع وقوة العلاقة بين متغيرين، ويرمز له في حالة المجتمع بالرمز ρ (وتقرأ "رُو")، وفي حالة العينة بالرمز r ، وحيث أننا في كثير من النواحي التطبيقية نتعامل مع بيانات عينة مسحوبة من المجتمع، سوف نهتم بحساب معامل الارتباط في العينة r كتقدير لمعامل الارتباط في المجتمع، ومن التحديد السابق للغرض من معامل الارتباط، نجد أنه يركز على نقطتين هما:

- نوع العلاقة:ـ وتأخذ ثلاث أنواع حسب إشارة معامل الارتباط كما يلي:
- ✓ إذا كانت إشارة معامل الارتباط سالبة ($r < 0$) توجد علاقة عكسية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه انخفاض في المتغير الثاني، والعكس.
- ✓ إذا كانت إشارة معامل الارتباط موجبة ($r > 0$) توجد علاقة طردية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني، والعكس .
- ✓ إذا كان قيمة معامل الارتباط صفرا ($r = 0$) دل ذلك على انعدام العلاقة بين

المتغيرين.

قوة العلاقة:ـ

يمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن $[\pm 1]$، حيث أن قيمة معامل الارتباط تقع في المدى $[-1 \leq r \leq 1]$.

⌖ **معامل الارتباط الخطى البسيط " لبيرسون" Karl Pearson**

في حالة جمع بيانات عن متغيرين كميين $(y\ ,\ x)$، يمكن قياس الارتباط بينهما، باستخدام طريقة "بيرسون" *Pearson*، ومن الأمثلة على ذلك: قياس العلاقة بين الوزن والطول، والعلاقة بين الإنتاج والتكلفة، والعلاقة بين الإنفاق الاستهلاكي والدخل، والعلاقة بين الدرجة التي حصل عليها الطالب وعدد ساعات الاستذكار، وهكذا الأمثلة على ذلك كثيرة.

إذا افترضنا أن لدينا عينة مكونة من n من أزواج المشاهدات $(x_1\ ,\ y_1),\ (x_2,\ y_2),\ ....,(x_n\ ,\ y_n)$ لمتغيرين X و Y، فإن معامل الارتباط لبيرسون يعرف على أنه:

$$r = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sqrt{\left[\sum_{i=1}^{n} X_i^{\,2} - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}\right]\left[\sum_{i=1}^{n} Y_i^2 - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n}\right]}} \qquad (1\text{-}9)$$

فإذا نظرنا إلى المعادلة السابقة عرفنا مقدار المعاناة عند حساب مقدار معامل الارتباط ناهيك عن حجم التحليل والحسابات إذا زادت حجم العينة و الآن نتعرف كيف يمكن تطبيق ذلك بمنتهى السهولة باستخدام البرنامج وذلك من خلال المثال القادم:

مثال(9-1):

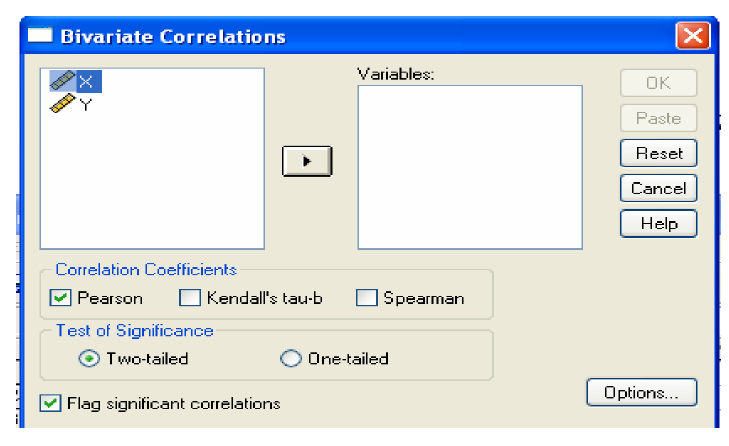
احسب معامل الارتباط المناسب بين درجة الطالب في الرياضيات ودرجته في الإحصاء باستخدام البيانات الآتية لعينة من خمسة طلاب:

جدول (1-9)

درجة الرياضة (x)	15	19	11	3	7
درجة الإحصاء (y)	18	20	16	14	12

الحل:

من شريط قوائم نختار Analyze ثم نختار Correlate ثم نختار Bivariate فيظهر الشكل الآتي:



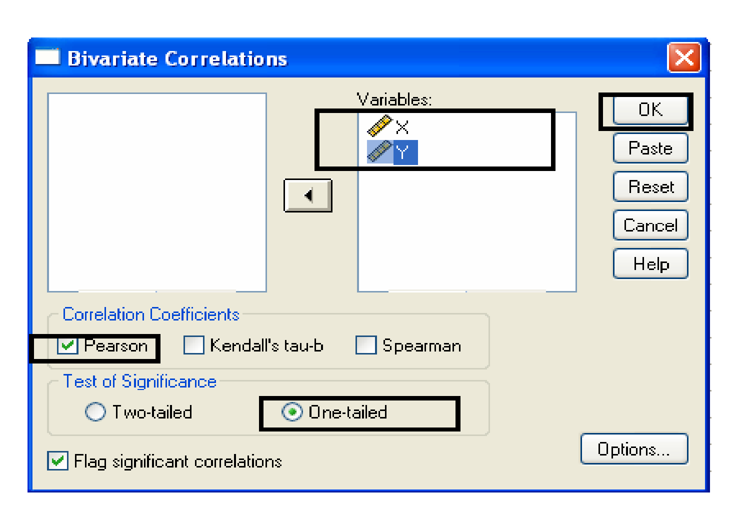
شكل (1-9)

ونرى أن الشكل السابق يتضمن مجموعة من الأجزاء:

- عمود المتغيرات المتاحة وهو العمود الأول من جهة اليسار ومنه يتم نقل المتغيرات إلى العمود الذي بجواره لنبدأ حساب معامل الارتباط وهنا نقوم بنقل X، Y.
- Correlation Coefficients: وهو يتيح لنا تحديد نوع معامل الارتباط الذي نريده ونلاحظ أنه متوفر لدينا 3 اختيارات:
  - ✓ Pearson: وهو معامل ارتباط الظواهر الكمية سواء كانت مبوبة أو غير مبوبة.
  - ✓ Kendall's tau: وهو معامل ارتباط الظواهر الترتيبية ويفضل للعينات الصغيرة والبيانات الغير مبوبة .

- ✓ Spearman: وهـو أيضـا معامـل الارتبـاط بـين الظـواهر الكميـة والترتيبـة ويفضل للعينات الصغيرة ويستخدم في حالة البيانات الغير مبوبة.
- ✓ وهنا نقوم بتحديد معامل بيرسون.
- ✓ Test of Siginificance: ولأننا نستخدم بيانات عينة فإننا نحتاج إلى التأكد من قيمة معامل الارتباط فنحتاج إلى اختبار هذة القيمة وهنا نختار إما الاختبار من جهة أو جهتين.
- ✓ Flag Significant Correlation: هنا تستخدم لتعليم الارتباط بعلامة نجمة (star) في حالة الارتباط المعنوي (الدال) أي الارتباط الحقيقي والذي يؤكده الاختبار.

وبعد الاختيار السابق يظهر الشكل (9-3) في الشكل الآتي:



شكل (9-2)

⌖ ثم ننقر على ok فيظهر الجدول الآتي:

جدول (9-2)

Correlations

		X	Y
X	Pearson Correlation	1	.900*
	Sig. (1-tailed)		.019
	N	5	5
Y	Pearson Correlation	.900*	1
	Sig. (1-tailed)	.019	
	N	5	5

*. Correlation is significant at the 0.05 level (1-tailed).

- ونجد أن معامل الارتباط بين درجة الطالب في الإحصاء ودرجته في الرياضيات مقدارها 0.9 وهذا يعني أن الارتباط طردي وقوي جدا ولاختبار الفرض الصفري:

$$H_0 : \rho = 0$$

ضد الفرض البديل:

$$H_A : \rho \neq 0.$$

ونلاحظ أن الفرض البديل لا يساوي صفر وهذا ما دل عليه اختيارنا السابق أن الاختبار ذو اتجاهين فإذا افترضنا أننا نريد اختبار ذلك بمستوى معنوية مقدارها $\alpha = 0.05$ فإن قيمة

p-value $= 0.019 < 0.05 = \alpha$ دل ذلك على أن الارتباط بين درجة الرياضيات و الإحصاء ارتباط يختلف عن الصفر بمستوى معنوي مقداره 0.05 أي أننا نقبل الفرض البديل القائل بأن الارتباط لا يساوي صفر.

- **معامل ارتباط الرتب (لسبيرمان) Spearman**

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتيبين، ومثال على ذلك قياس العلاقة بين تقديرات الطلبة في مادتين ، أو العلاقة بين درجة تفضيل المستهلك لسلعة معينة ، ومستوى الدخل، فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية ، ويطلق على هذا المعامل " معامل ارتباط سبيرمان " Spearman.

مثـــال (9-2):

فيما يلي بيانات درجة تفضيل ومستوى دخل لعشرة مستهلكين لسلعة ما كعينة والمطلوب معرفة الارتباط بين مستوى الدخل ودرجة تفضيل المستهلكين للسلعة وما مدلول ذلك بالنسبة لجميع المستهلكين:

جدول (9-3)

مستوى التفضيل	أفضل	لا أفضل	أفضل	لا أفضل	لا أفضل	لا أفضل	أفضل	لا أفضل	أفضل	أفضل
مقدار الدخل الشهري	100	200	250	400	130	270	700	500	330	200

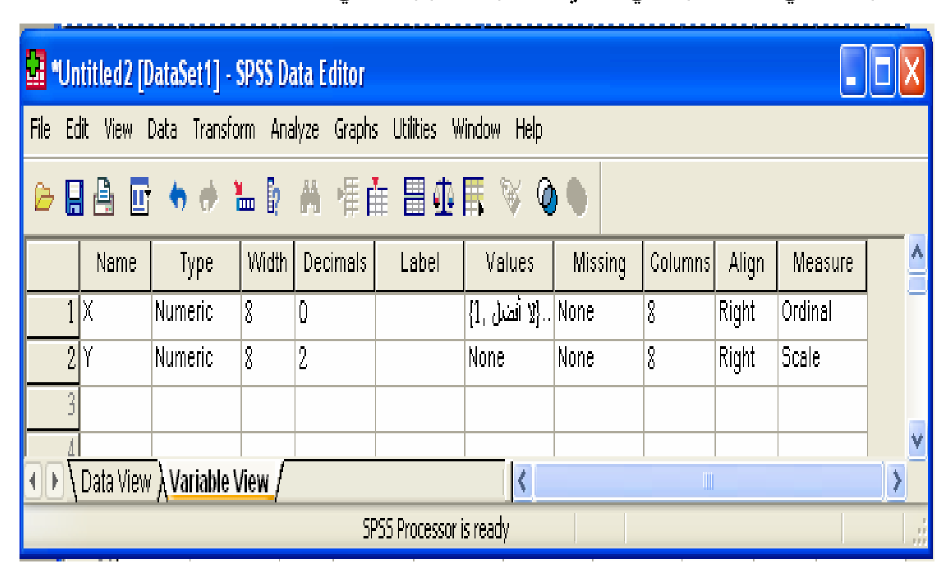
الحــل:

خلال البرنامج نقوم بإدخال البيانات على الفرض أن "أفضل" نرمز لها بالرمز 1 و"لا أفضل" برمز 2.

ونقوم بتعريف المتغير الترتيبي x كما يلي في محرر المتغيرات نقوم بتعريف قيم المتغير في عمود values كما يلي:

شكل (9-3)

ثم في عمود measure نعرف المتغير X على أنه متغير ترتيبي ordinal ونعرف المتغير Y علي أنه متغير كمي عادي فيظهر المحرر كالآتي:ْ



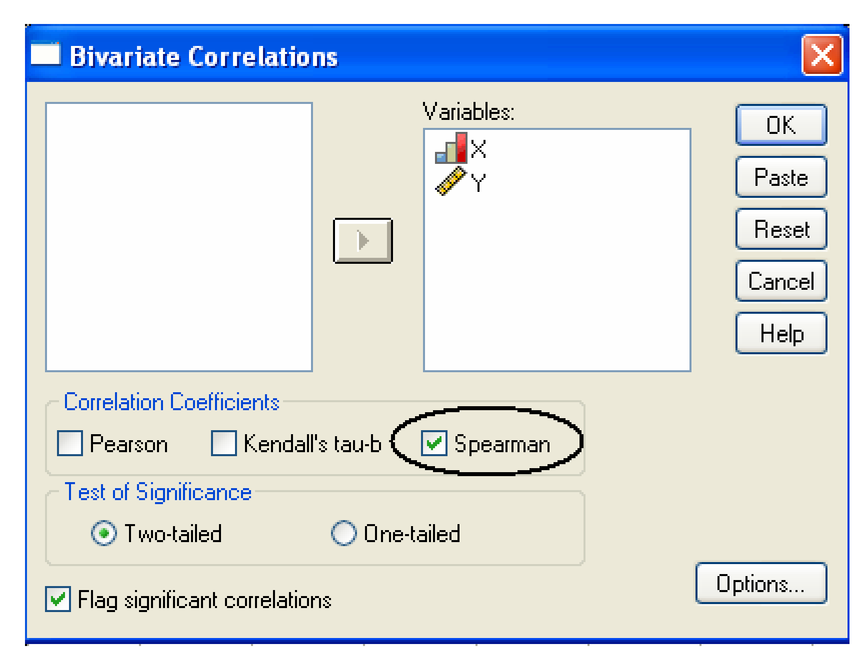
شكل (9-4)

ثم شاشة محرر البيانات نقوم بإدخال البيانات السابقة فتظهر في الشكل الآتي:

	X	Y
1	1	100.00
2	2	200.00
3	1	250.00
4	2	400.00
5	2	130.00
6	2	270.00
7	1	700.00
8	2	500.00
9	1	330.00
10	1	200.00

شكل (9-5)

من شريط قوائم نختار Analyze ثم نختار Correlate ثم نختار Bivariate فيظهر الشكل الآتي:



شكل (6-9)

ونلاحظ أن المتغير x معرف في عمود المتغير المستخدمة في التحليل (العمود الأول من اليمين) وأننا اختارنا spearman كمعامل للارتباط ونفس الاختيارات السابقة في مقياس Pearson ثم ننقر على ok فيظهر الجدول الآتي:

جدول (4-9)

Correlations

			X	Y
Spearman's rho	X	Correlation Coefficien	1.000	.070
		Sig. (2-tailed)	.	.848
		N	10	10
	Y	Correlation Coefficien	.070	1.000
		Sig. (2-tailed)	.848	.
		N	10	10

ونجد أن معامل الارتباط بين درجة تفضيل المستهلكين لسلعة ومستوى دخولهم مقداره 0.070 وهذا يعني أن الارتباط طردي وضعيف ولاختبار الفرض القائل:

$$H_0 : \rho = 0$$

ضد الفرض البديل القائل :

$$H_A : \rho \neq 0 .$$

ونلاحظ أن الفرض البديل معامل الارتباط لا يساوي صفر وهذا ما دل عليه اختيارنا السابق أن الاختبار ذو اتجاهين فإذا افترضنا أننا نريد اختبار ذلك بمستوى معنوية مقدارها $\alpha = 0.05$ فإن قيمة $\alpha = 0.05 < 0.424 =$ p-value دل ذلك على أن الارتباط بين تفضيل المستهلكين لسلعة ومستوى دخولهم ارتباط يساوي الصفر بمستوى معنوي مقداره 0.05 أي أننا نقبل الفرض العدم القائل الارتباط يساوي صفر في كامل بيانات المجتمع.

(3-9) الارتباط الجزئي Partial Correlation

يقيس معامل الارتباط الجزئي قوة العلاقة بين متغيرين بثبوت متغير ثالث أو أكثر. مثلا قد نحصل على قيمة عالية لمعامل الارتباط البسيط للعلاقة بين أسعار اللحوم البيضاء واللحوم الحمراء فقد لا توجد علاقة فعلية بين المتغيرين ولكن كلا المتغيرين يتأثر بمعامل ثالث هو المستوى العام للأسعار فإذا استبعدنا المستوى العام للأسعار (أو تثبيته) عند قياس العلاقة بين أسعار اللحوم البيضاء والحمراء فسيتم الحصول على قيمة أقل لمعامل الارتباط وهذا يعرف بالارتباط الجزئي . علما أنه يمكن استبعاد أي عدد من المتغيرات عند قياس العلاقة بين ظاهرتين.

والآتي مثال يوضح كيفية تطبيق ذلك من خلال البرنامج:

مثـــال (3-9):

البيانات الآتية تعطي الدخل الشهري لمجموعة أسر (X) والإنفاق الشهري على الطعام (Y) وحجم مدخراتها الشهرية (Z) وذلك في عينة من تسعة أسر:

جدول (9-5)

الأسرة	1	2	3	4	5	6	7	8	9
x	40	65	90	60	75	80	120	45	65
Y	22	28	40	20	32	32	50	20	35
Z	6	10	8	25	12	7	10	5	7

المطلوب:

1- حساب معامل الارتباط الجزئي لـ x و y بثوت z.

2- حساب معامل الارتباط الجزئي لـ x و z بثوت y.

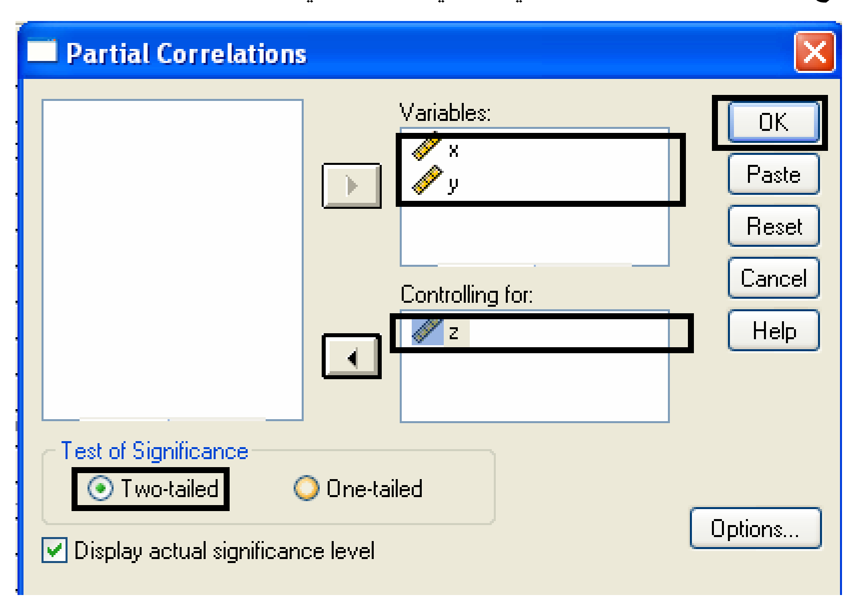
3- حساب معامل الارتباط الجزئي لـ z و y بثوت x.

4- اختبار معنوية (دلالة) الارتباط بمستوى معنوية قدره 0.05.

الحل:

أولا: حساب معامل الارتباط الجزئي لـ x و y بثوت z.

من شريط قوائم نختار Analyze ثم نختار Correlate ثم نختار Partial فيظهر مربع حوار معامل الارتباط الجزئي كما في الشكل التالي:



شكل (9-7)

في خانة variables يتم إدخال المتغيرات التي يراد حساب معامل الارتباط الجزئي لها x، y ، وفي خانة Controlling for يتم إدخال المتغير (المتغيرات) الذي يراد استبعاد أثره (z).

وعند النقر على زر ok نحصل على النتيجة التالية:

جدول (6-9)

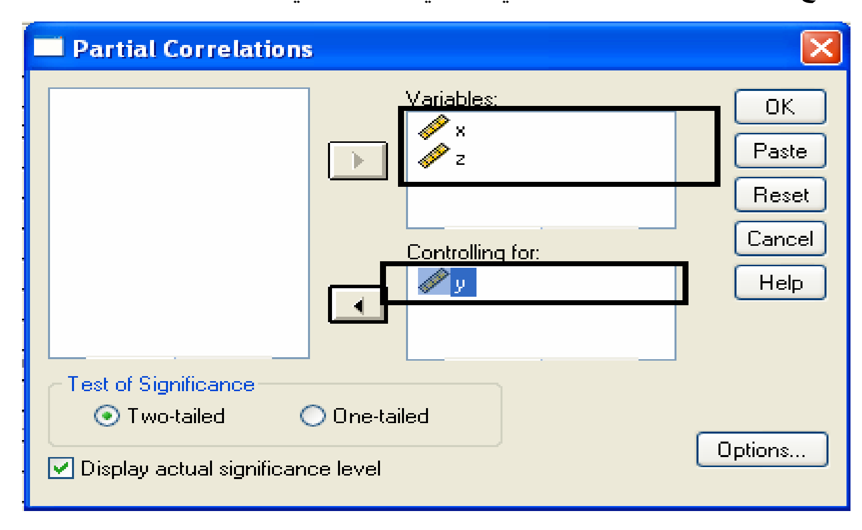
Correlations

Control Variables			x	y
z	x	Correlation	1.000	.965
		Significance (2-tailed)	.	.000
		df	0	6
	y	Correlation	.965	1.000
		Significance (2-tailed)	.000	.
		df	6	0

نلاحظ أن قيمة معامل الارتباط من الشكل السابق تساوي 0.965 أي أنه توجد علاقة طردية وقوية بين الدخل والإنفاق وكما أن قيمة p-value=0 وهي أقل من قيمة مستوى المعنوية مقسوم على 2=0.025 حيث أننا نختبر فرض عدم ذو اتجاهين ولذلك فإننا نقبل الفرض البديل القائل أن الارتباط بين المتغيرات في المجتمع لا يساوي صفر0

ثانياً: حساب معامل الارتباط الجزئي لـ x و z بثوت y.

من شريط قوائم نختار Analyze ثم نختار Correlate ثم نختار Partial فيظهر مربع حوار معامل الارتباط الجزئي كما في الشكل التالي:



شكل (8-9)

في خانة variables يتم إدخال المتغيرات (x، z) ، وفي خانة Controlling for يتم إدخال المتغير (y). عند النقر على زر ok نحصل على النتيجة التالية:

جدول (9-7)

Correlations

Control Variables			x	z
y	x	Correlation	1.000	.693
		Significance (2-tailed)	.	.057
		df	0	6
	z	Correlation	.693	1.000
		Significance (2-tailed)	.057	.
		df	6	0

إن قيمة معامل الارتباط من الشكل السابق تساوي 0.693 أي أنه توجد علاقة طردية بين الدخل الادخار وكما أن قيمة p-value=0.057 وهي أكبر من قيمة مستوى المعنوية مقسوم على 2 حيث أننا نختبر فرض عدم ذو اتجاهين ولذلك فإننا نقبل الفرض العدم القائل أن الارتباط بين المتغيرات في المجتمع يساوي صفر.

ثالثاً: حساب معامل الارتباط الجزئي لـ z و y بثوت x.

كما فعلنا في أولا وثانيا في إدخال وتعريف المتغيرات نفعل في تلك الخطوة وتظهر نتيجة معامل الارتباط كآلاتي:

جدول (9-8)

Correlations

Control Variables			z	y
x	z	Correlation	1.000	-.710
		Significance (2-tailed)	.	.048
		df	0	6
	y	Correlation	-.710	1.000
		Significance (2-tailed)	.048	.
		df	6	0

فنجد أن قيمة معامل الارتباط بين الاستهلاك والإدخار تساوي 0.710 - وهذا يعني أنه توجد علاقة بين الاستهلاك والإدخار عكسية وقوية، ونجد أيضا أن قيمة p-value=0.048 وهي أكبر من قيمة مستوى المعنوية مقسوم على 2=0.025 حيث أننا نختبر فرض عدم ذو اتجاهين ولذلك فإننا نقبل الفرض العدم القائل أن الارتباط بين المتغيرات في المجتمع يساوي صفر في ضوء بيانات العينة.

الفصـــل العــاشـــر

تحـــليـــل الانحـــدار

Regression Analysis

(1-10) مقدمة:

إن نموذج الانحدار يعبر عن علاقة بين متغير تابع وبين واحد أو أكثر من المتغيرات المستقلة فإذا احتوى النموذج على متغير مستقل واحد فيعرف بنموذج الانحدار البسيط Model Simple Regression أو ما يسمى Ordinary Least Squares و إذا احتوى على أكثر من متغير مستقل فهو نموذج الانحدار المتعدد كما أن النموذج قد يكون خطياً أو غير خطي .

(2-10) نموذج الانحدار الخطي البسيط Simple Regression

إن الغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر، ومن الأمثلة على ذلك ما يلي:

- دراسة أثر كمية السماد على إنتاجية القمح.
- دراسة أثر الإنتاج على التكلفة.
- دراسة أثر كمية البروتين التي يتناولها الأبقار على الزيادة في الوزن.
- أثر الدخل على الإنفاق الاستهلاكي.
- دراسة أثر الذكاء على التحصيل الدراسي.

وهكذا هناك أمثلة في كثير من النواحي الاقتصادية، والزراعية، والتجارية، والعلوم السلوكية، وغيرها من المجالات الأخرى.

في تحليل الانحدار البسيط، نجد أن الباحث يهتم بدراسة أثر أحد المتغيرين ويسمى بالمتغير المستقل أو المتنبأ منه، على المتغير الثاني ويسمى بالمتغير التابع أو المتنبأ به، ومن ثم يمكن عرض نموذج الانحدار الخطي في شكل معادلة خطية من الدرجة الأولى، تعكس المتغير التابع كدالة في المتغير المستقل كما يلي:

$$y = \beta_0 + \beta_1 x + e \qquad (1\text{-}10)$$

حيث أن:

- ✓ y: هو المتغير التابع (الذي يتأثر)
- ✓ x: هو المتغير المستقل (الذي يؤثر)
- ✓ β_0: هو الجزء المقطوع من المحور الرأسي y ، وهو يعكس قيمة المتغير التابع في حالة انعدام قيمة المتغير المستقل x ، أي في حالة $x=0$
- ✓ β_1: ميل الخط المستقيم $(\beta_0+\beta_1 x)$ ، ويعكس مقدار التغير في y إذا تغيرت x بوحدة واحدة.
- ✓ e: هو الخطأ العشوائي، والذي يعبر عن الفرق بين القيمة الفعلية y ، والقيمة المقدرة $\hat{y}=\beta_0+\beta_1 x$ ، أي أن : $e=y-(\beta_0+\beta_1 x)$.

⌖ تقدير نموذج الانحدار الخطي البسيط

يمكن تقدير معاملات الانحدار $(\beta_1 \ , \ \beta_0)$ في النموذج (10-1) باستخدام طريقة المربعات الصغرى، وهذا التقدير هو الذي يجعل مجموع مربعات الأخطاء العشوائية $\sum e^2=\sum(y-(\beta_0+\beta_1 x))^2$ أقل ما يمكن، ويحسب هذا التقدير بالمعادلة التالية:

$$\hat{\beta}_1=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-\left(\sum x\right)^2} \qquad (2\text{-}10)$$

حيث أن $\bar{x}$ هو الوسط الحسابي لقيم x ، $\bar{y}$ هو الوسط الحسابي لقيم y ، وتكون القيمة المقدرة للمتغير التابع هو: $\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x$ ، ويطلق على هذا التقدير" تقدير معادلة انحدار y على x.

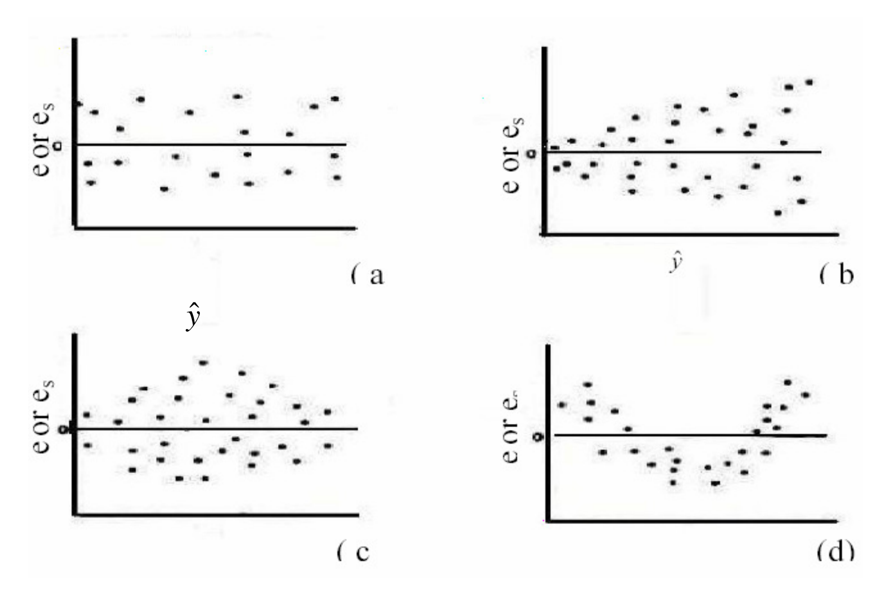
علماً بان نموذج الانحدار الخطي البسيط يجب أن يحقق مجموعة الفروض الآتية:

- ⌖ وجود علاقة خطية بين X و Y.
- ⌖ أن الأخطاء العشوائية تتوزع بمتوسط مساوي للصفر.
- ⌖ أن الأخطاء العشوائية لها تباين ثابت يساوي σ^2 (فرضية تجانس تباين الخطأ العشوائية Homoscedasticity).

- الأخطاء تتوزع طبيعياً وهذا الشرط ليس ضرورياً لتقدير المعالم بطريقة OLS ولكنه ضروري لاختبار الفرضيات المعلقة بمعاملات الانحدار β_0 و β_1.

عدم وجود ارتباط ذاتي Autocorrelation بين الأخطاء العشوائية.

ويمكن التحقق من توفر فرضيات النموذج الخطي البسيط من خلال تخطيط Scatter plots بتمثيل $\hat{y}$ على المحور الأفقي (أو X) يقابله الخطأ العشوائي e أو الأخطاء المعيارية Standardized Residual يرمز لها e_s على المحور الرأسي. كما هو واضح في الشكل الآتي:



(a) توفر فروض التحليل جميعها (عدم وجود مشكلة).

(b) زيادة تباين الخطأ العشوائي بزيادة $\hat{y}$.

(c) زيادة وتناقص في تباين الخطأ العشوائي (مشكلة عدم تجانس تباين الخطأ العشوائي).

(d) عدم ملائمة العلاقة الخطية (يتوجب استعمال نماذج أخرى مثلا نموذج الدرجة الثانية).

وفيما يلي نبين كيف يمكننا اختبار هذه الشروط وكيفية تقدير معادلة الانحدار وتحديد الأخطاء باستعمال البرنامج:

تطبيقات:

مثـال (10-1)

فيمـا يلـي بيانـات عـن كميـة البـروتين اليـومي بـالجرام التـي يحتاجهـا العجـل الرضـيع، ومقـدار الزيـادة فـي وزن العجـل بـالكليو جـرام، وذلـك لعينـة مـن العجـول الرضيعة حجمها 10.

جدول (10-1)

Xكمية البروتين	10	11	14	15	20	25	46	50	59	70
Yالزيادة في الوزن	10	10	12	12	13	13	19	15	16	20

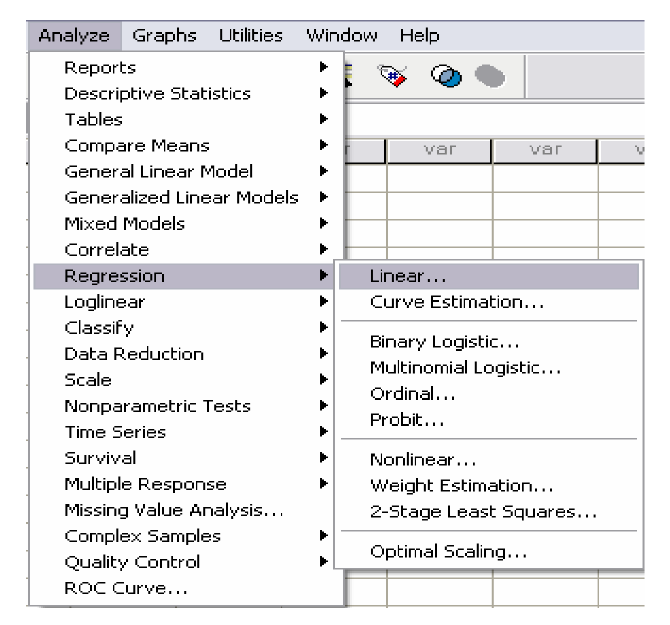
والمطلوب :

- قدر معادلة انحدار الوزن y على كمية البروتين x.
- فسر معادلة الانحدار.
- ما هو مقدار الزيادة في الوزن عند إعطاء العجل 50 جرام من البروتين ؟
- وما هو مقدار الخطأ العشوائي؟
- استخراج فترة ثقة 95% لكل من معلمتي الانحدار β_0 و β_1 .
- استخراج جدول تحليل التباين ANOVA للمعاملات.
- اختبر جودة توفيق النموذج الخطـي (باستعمال معامل التحديد R^2) مع تحليل الأخطاء العشوائية بالرسم البياني.
- اختبر التوزيع الطبيعي للأخطاء العشوائية بيانياً.

الحل:

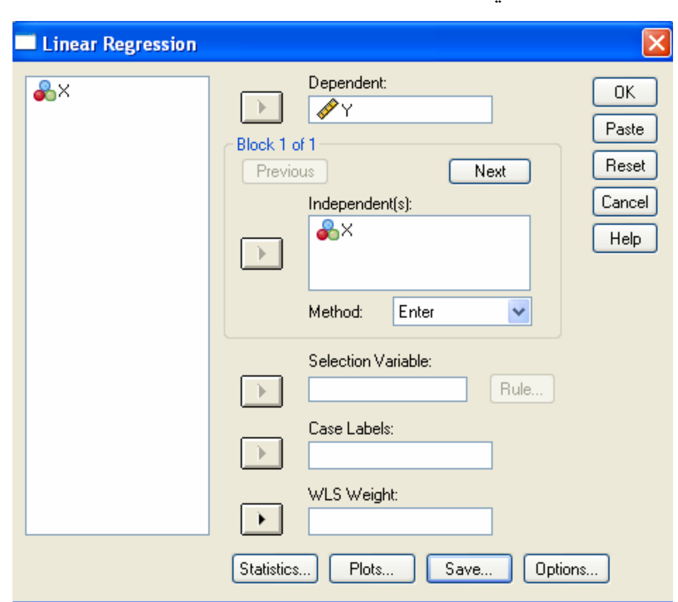
لتنفيذ المطلوب السابق باستخدام البرنامج نتبع الخطوات التالية:

Analyze ثم $\Leftarrow$ Regression $\Leftarrow$ Linear كما في الشكل (10-1):



شكل (1-10)

فيظهر صندوق الحوار التالي:



شكل (2-10)

ونجد أن الشكل (2-10) يتكون من:

Dependent: يمثل المتغير التابع.

Independent: وهو يمثل المتغير المستقل ويمكننا إدخال مجموعة من المتغيرات المستقلة هنا كما سوف نرى في مبحث الانحدار المتعدد ولأننا نريد معادلة انحدار ل y على X فإننا نعني أن y متغير تابع X متغير مستقل .

Method: نوع الطريقة المستخدمة في الانحدار (الطريقة الاعتيادية هي Enter).

Selection Variable: يستعمل في تحديد التحليل لمجموعة معينة من الحالات التي لها قيمة معينة لمتغير الاختيار (مثلا اقتصار نموذج الانحدار على الحالات التي تكون فيها قيمة المتغير Observat أكبر من 5) يتم التحديد بواسطة الزر Rule.

Case Lebels: متغير يستخدم قيمة كعناوين لنقاط شكل الانتشار Scatterplots.

Statistics: عند النقر على هذا الزر يظهر مربع الحوار الآتي:

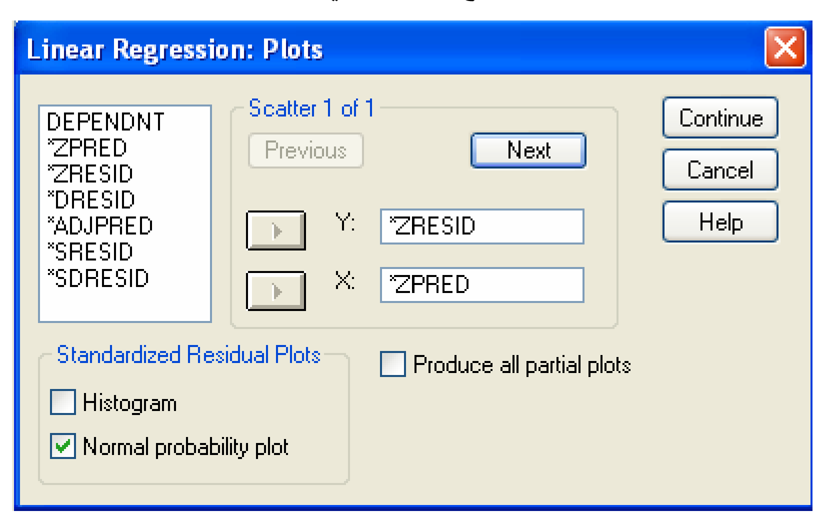
شكل (3-10)

وقد تم تأشير الخيارات التالية:

- ✓ Estimate: لتقدير معالم نموذج الانحدار واختبارات t المرافقة .
- ✓ Confidence Interval: لتقدير فترة ثقة %95 لكل من معلمتي الانحدار.
- ✓ Model Fit: لعرض R^2 و ANOVA .

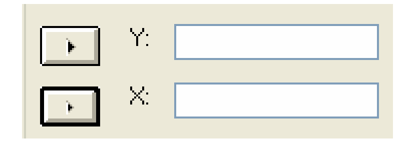
وبعد الانتهاء النقر على زر Continue. وهناك زر آخر في شكل (2-10) بعنوان

وعند النقر عليه يظهر مربع الحوار الآتي: Plots



شكل (4-10)

ونلاحظ أننا اخترنا Normal probability plot: لاختبار التوزيع الطبيعي للأخطاء العشوائية المعيارية . كما أننا نجد أعلى شكل (4-10) الشكل الآتي:



في هذا الشكل يتم ادخال المتغيرات التي نود أن تظهر في شكل تحليل الاخطاء في المحور السيني والمحور الصادي والمتغيرات موجود في الجانب الأيمن من الشكل السابق ويوجد عدة خيارات:

1- Standardized predicted values (*ZPRED): القيمة المعيارية لقيمه المتوقعة $\hat{y}$.

2- Standardized residuals (*ZRESID): القيمة المعيارية للاخطاء.

3- Deleted residuals (*DRESID): حذف الاخطاء.

4- Adjusted predicted values (*ADJPRED): القيمة المعيارية المعدلة لقيمة متوقعة $\hat{y}$.

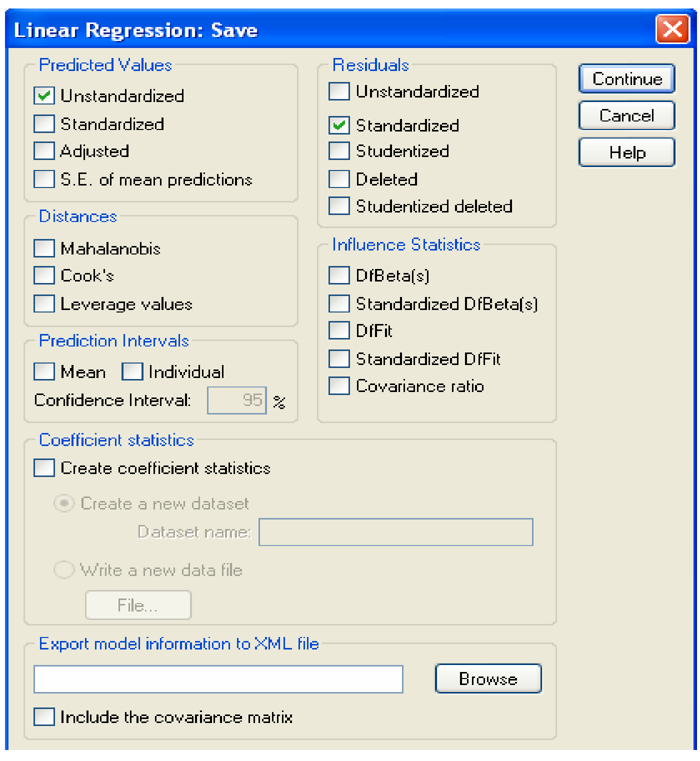
5- Studentized residuals (*SRESID): جعل الاخطاء تتبع توزيع t.

6- Studentized deleted residuals (*SDRESID): الاخطاء المحذوفة

والتي تتبع توزيع t.

وفي تلك الحالة يتم تخصص كلا من :ZPRED لمحور السينات ،ZRESID لمحور الصادات وبعد الانتهاء يتم النقر على زر Continues.

وهناك زر آخر في شكل (2-10) بعنوان Save وعند النقر عليه يظهر مربع الحوار الآتي:



شكل (5-10)

ونلاحظ أننا اختارنا Values predicted Unstandardized أي $\hat{y}$ وكذلك Residual Standardized أي e_s سيتم استعمال هذين المتغيرين لرسم Scatterplots الجزء الثاني من المطلوب م ، علماً أنه سيتم إضافة هذين المتغيرين إلى ورقة Editor Data إلى جانب متغيرات X و Y، Observat حيث يضاف المتغير Values predicted Unstandardized باسم Pre_1 ويضاف Residual Standardized باسم Zre_1. وبعد الانتهاء النقر على زر Continues.

و في النهاية النقر على Ok في الشكل (2-10) لإظهار النتائج كالآتي:

جدول (2-10)

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B Lower Bound	95% Confidence Interval for B Upper Bound
1	(Constant)	9.436	.864		10.926	.000	7.444	11.427
	X	.143	.023	.913	6.325	.000	.091	.195

a. Dependent Variable: Y

من خلال الجدول أعلاه يمكن كتابة النموذج كما يلي:

$$\hat{y} = 9.436 + 0.143x$$

وتشير المعادلة السابقة بأن زيادة البروتين مقدار جرام يعمل على زيادة الوزن بمقدار 0.143 جرام بينما توجد زيادة ليست متعلقة بالبرويتن مقدارها 09.436 لكل جرام في الوزن وعليه فإن زيادة البروتين بمقدار 50 جرام يعمل على أن يكون

$$\hat{y} = 9.436 + 0.143(50)$$
$$= 16.436$$

ومن ثم مقدار الخطأ العشوائي يساوي:

$$\hat{e}_{x=50} = y_{x=50} - \hat{y}_{x=50}$$
$$= 15 - 16.43$$
$$= -1.43$$

ويستعمل اختبار t لاختبار الفرضية التالية لمعلمة الميل β_1:

$$H_0 : \beta_1 = 0$$
$$H_A : \beta_1 \neq 0$$

ويستعمل اختبار t لاختبار الفرضية التالية لمعلمة الميل β_0:

$$H_0 : \beta_0 = 0$$
$$H_A : \beta_0 \neq 0$$

نستخدم قيمة *P-value* المرافقة لإحصائية t للمعلمة في الاختبار كما يلي:

إذا كانت *P-value<0.05* نرفض فرضية العدم بمستوى *5%*.

إذا كانت *P-value<0.01* نقبل فرضية العدم بمستوى *1%*.

عكس هذا نقبل فرضية العدم.

ونلاحظ أن *P-value* لمعلمة الميل تساوي صفر وهي أقل من 0.01 ، إن *P-value* لمعلمة الحد الثابت تساوي صفر وهي أقل من 0.01 ولهذا نرفض فرضي العدم لكل من المعلمتين أي أن كلا من المعلمتين تختلف جوهرياً عن الصفر. إن ظهور معلمة الميل معنوية يعكس أهمية متغير البروتين في التأثير علي الوزن في النموذج.

ويمكن كتابة فترة الثقة للحد الثابت وذلك باستعمال نتائج الشكل (10-6) كما يلي:

$$\Pr[7.444 \le \beta_0 \le 11.427] = 95\%$$

حيث أن Pr تمثل الاحتمال كما يمكن كتابة فترة ثقة 95% لمعلمة الميل كما يلي:

$$\Pr[0.091 \le \beta_0 \le 0.195] = 95\%$$

الجدول التالي يعرف بجدول تحليل التباين ANOVA و يشتمل على إحصائية F لاختبار نفس الفرضية الخاصة بمعلمة الميل β_1 وهذا الاختبار مكافئ تماماً لاختبار t لمعلمة الميل (لاحظ أن قيمة P-value متساوية لكلا الاختبارين) علما أن $F = t^2$.

جدول (10-3)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	90.000	1	90.000	40.001	.000[a]
	Residual	18.000	8	2.250		
	Total	108.000	9			

a. Predictors: (Constant), X

b. Dependent Variable: Y

الجدول التالي يتضمن أهم مؤشر لنموذج الانحدار وهو معامل التحديد Coefficient of Determination ويرمز له R^2 ويعتبر مقياساً لجودة توفيق النموذج :

جدول (4-10)

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.913[a]	.833	.813	1.49999

a. Predictors: (Constant), X

b. Dependent Variable: Y

ويحتسب من جدول تحليل التباين كما يلي:

$$R^2 = \frac{\text{Explained Variations}}{\text{Total Variations}}$$
$$= \frac{SSR}{SST}$$
$$= \frac{90}{108.00}$$
$$= 0.833333 \quad , \quad 0 \le R^2 \le 1$$

وتفسير ذلك أن 83.333% من التباينات (الانحرافات الكلية من قيم المتغير y) تفسرها العلاقة الخطية أي نموذج الانحدار وأن 15.76% من التباينات ترجع إلى عوامل عشوائية كأن تكون هناك متغيرات مهمة لم تتضمنها النموذج . وعلى العموم كلما اقتربت قيمة R^2 من 100% دل ذلك على جودة توفيق النموذج . هذا ، أن $r = \sqrt{R^2} = 0.91287$ حيث أن r معامل الارتباط الخطي البسيط لبيرسون وأن إشارة r هي نفس إشارة معلمة الميل أي أن الإشارة موجبة وهذا يعني أن العلاقة بين مقدار البروتين والوزن علاقة طردية وقوية.

يتصف معامل التحديد بأنه لو أضيف متغير مستقل للنموذج فإن قيمته سترتفع حتى لو لم تكن هناك أهمية للمتغير المستقل في النموذج حيث إن إضافة متغير مستقل إلى نموذج الانحدار يؤدي إلى زيادة R^2 بسبب زيادة مجموع المربعات العائدة للانحدار SSR مع ثبات مجموع المربعات الكلية SST ولهذا يتم احتساب معامل التحديد المصحح Adjusted R Square الذي يأخذ في الاعتبار النقصان الحاصل في درجات الحرية وقيمته دائما أقل من قيمة معامل التحديد (الغير المصحح) وفي هذا المثال

81.3% ولهذا يمكن القول أن النموذج جيد التوفيق.

أما الخطأ المعياري للتقدير Standard Error of Estimate فيقيس تشتت القيم المشاهدة عن خط الانحدار وأن الحصول على قيمة صغيرة لهذا المؤشر يعني صغر الأخطاء العشوائية وبالتالي جودة تمثيل خط الانحدار لنقاط شكل الانتشار.

لتحليل الأخطاء العشوائية بيانياً نكون شكل الانتشار Scatterplots بتمثيل القيم التقديرية $\hat{y}$ على المحور الأفقي و الأخطاء المعيارية على المحور العمودي أو الراسي وكما يلي :

ففي شاشة Data Editor نجد أنه كما سبق أن أوضحنا أنه سوف يتم إضافة المتغير Unstandardized predicted Values باسم Pre_1 ويضاف Standardized Residual باسم Zre_1 كما يلي:

	X	Y	PRE_1	ZRE_1
1	10	10.00	10.86212	-.57475
2	11	10.00	11.00475	-.66984
3	14	12.00	11.43264	.37824
4	15	12.00	11.57527	.28315
5	20	13.00	12.28843	.47439
6	25	13.00	13.00158	-.00105
7	46	19.00	15.99684	2.00212
8	50	15.00	16.56736	-1.04491
9	59	16.00	17.85104	-1.23403
10	70	20.00	19.41998	.38668

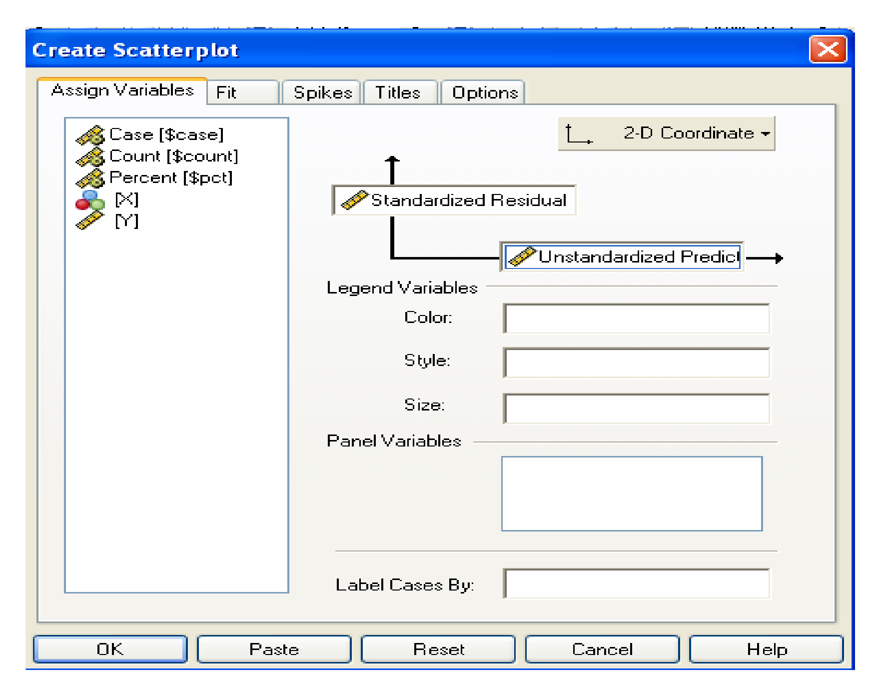
شكل (6-10)

ثم بعد ذلك من شريط القوائم نختار Graphs ⇐ Interactive ⇐ Scatter ⇐ Simple فيظهر صندوق حوار بعنوان Create Scatterplots :

فنقوم بنقر على Zre_1 لادخاله إلى الخانة Y في صندوق الحوار وبالمثل

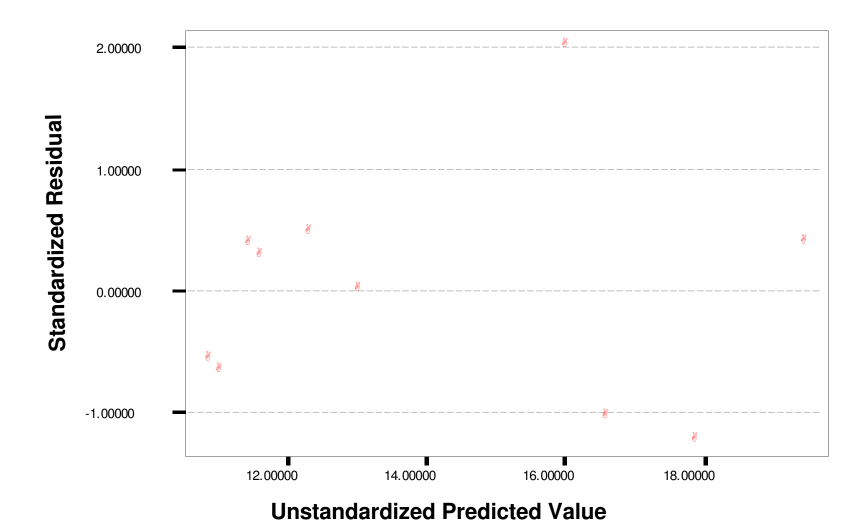
فنقوم بنقر على Pre_1 لادخاله إلى الخانة X في صندوق الحوار وبالمثل

فيظهر مربع الحوار Create Scatterplots كما يلي:



شكل (7-10)

فننقر على Ok فيظهر المخطط الآتي:



شكل (8-10)

نلاحظ أن النقاط تتوزع بشكل شريط أفقي متساوي حول الصفر مما يدل على توفر فرضيات التحليل بصورة عامة حيث لا يعاني النموذج من مشكلة عدم تجانس تباين الخطأ العشوائي و لا توجد حاجة لاستخدام علاقة من درجات أعلى.

مثال (10-2)

لدينا شركة تقوم بصناعة قطع غيار معينة على دفعات شهرية وقد أعطتنا الشركة البيانات لعشره دفعات في الجدول التالي:

جدول (10-5)

دورة الإنتاج	حجم الدفعة	ساعات العمل
1	30	73
2	20	50
3	60	128
4	80	170
5	40	87
6	50	108
7	60	135
8	30	69
9	70	148
10	60	132

تريد الشركة معرفة نموذج الانحدار الخطي الخاص لها لكي تتنبأ بأحجام الدفعات المستقبلية بناءا على ساعات العمل عند مستوى معنوية 0.05.

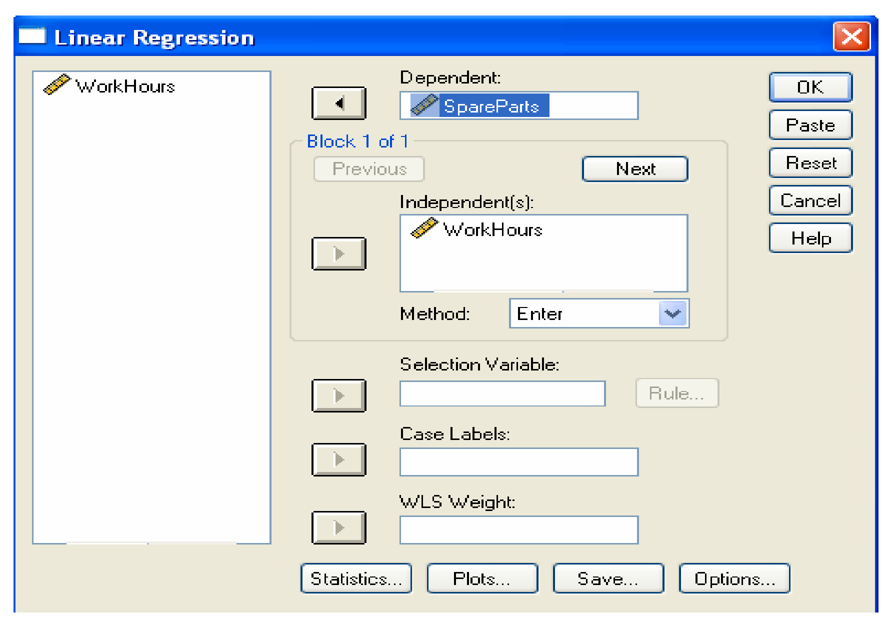
الحل:

ندخل البيانات في عمودين أحدهما لحجم الدفعة من قطع الغيار Spare Parts والآخر لساعات العمل على كل دفعة Work Hours كما يلي:

	SpareParts	WorkHours
1	30.00	73.00
2	20.00	50.00
3	60.00	128.00
4	80.00	170.00
5	40.00	87.00
6	50.00	108.00
7	60.00	135.00
8	30.00	69.00
9	70.00	148.00
10	60.00	132.00

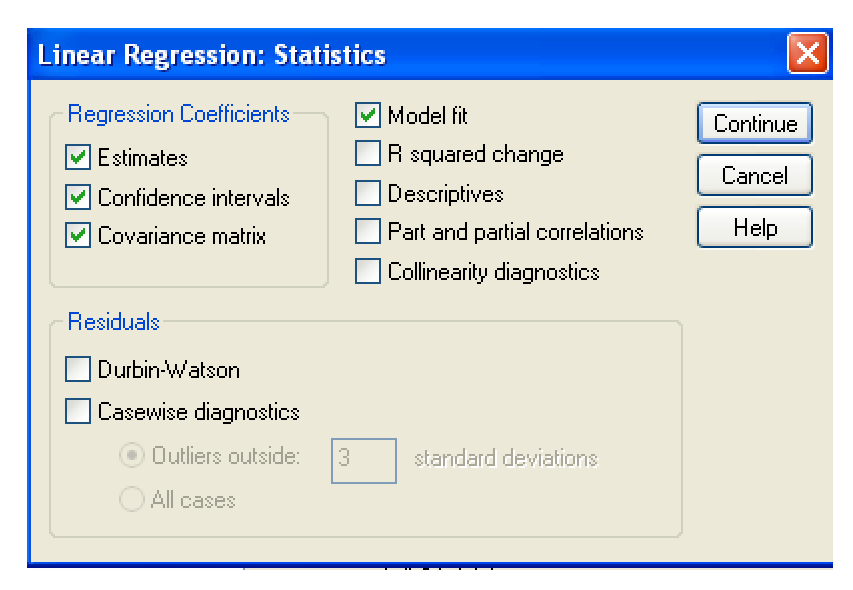
الشكل (10-9)

نقوم بتفعيل نافذة الانحدار ونختار المتغير المستقل هو Work Hours أما التابع فهو Spare Parts كما في الشكل التالي:



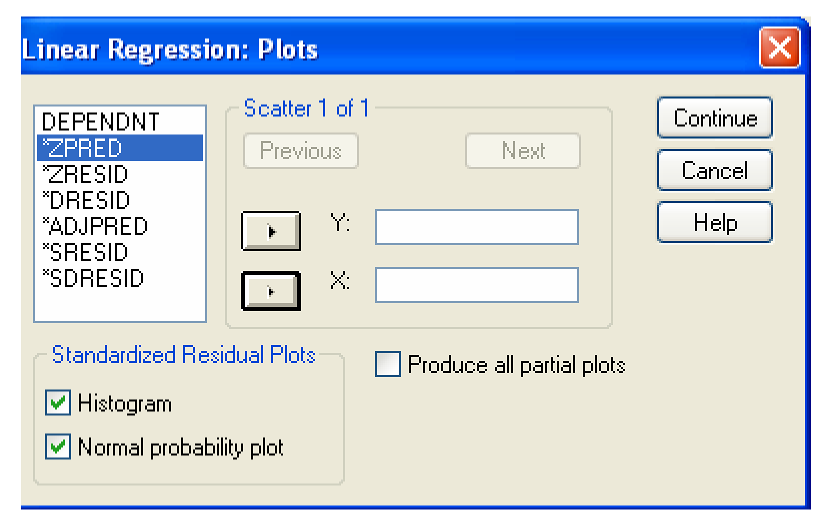
الشكل (10-10)

نختار الطريقة المعيارية وهي Enter في قائمة الخيارات Method لأنها الطريقة الخاصة بتقدير معاملات الانحدار المجهولة للمتغيرات المستقلة، بعد ذلك نقوم باختيار الأمر الفرعي Statistics و نتأكد من اختيار كل من Estimates، Confidence intervals،Model fit Covariance matrix ، كما في الشكل التالي:



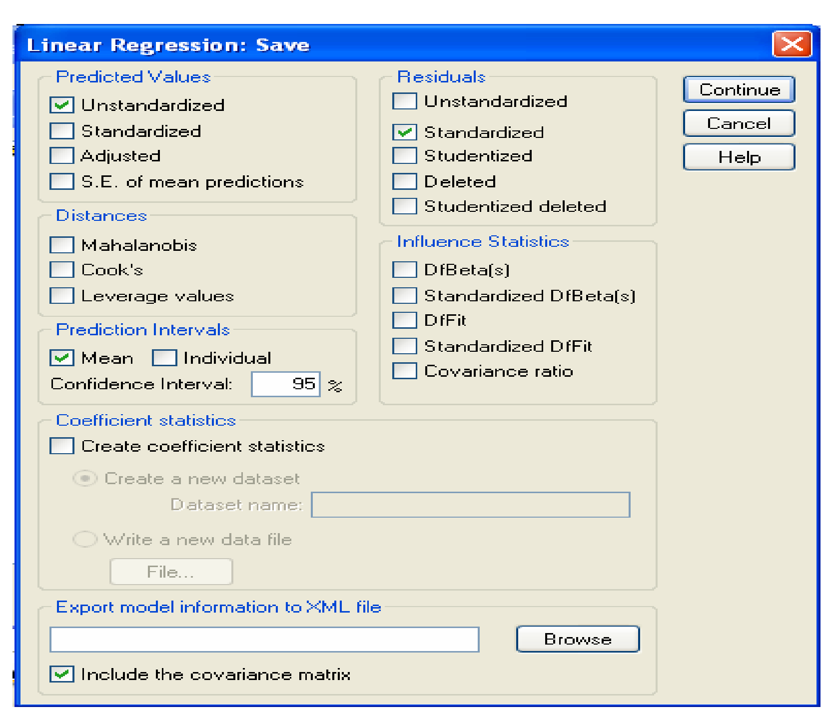
الشكل (10-11)

ثم نعود إلى مربع الحوار السابق بالنقر على Continue ونذهب إلى الأمر الفرعي Plots و نتأكد من اختيار Histogram و Normal probability plot كما يلي:



الشكل (10-12)

ثم نعود إلى مربع الحوار السابق بالنقر على Continue ونذهب إلى الأمر الفرعي Save ونتأكد من الاختيارات السابقة في المثال السابق كما يلي:



الشكل (10-13)

ثم نعود إلى مربع الحوار السابق بالنقر على Continue ثم ننقر OK فتظهر لدينا النتائج التالية:

يظهر لدينا جدول يبين الطريقة المستخدمة في الانحدار والمتغير التابع وهو Spare Parts والمتغيرات المستقلة الداخلة في نموذج الانحدار الخطي وفي حالتنا يوجد متغير واحد فقط وهو Work Hours كما يلي:

جدول (10-6)

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	WorkHours[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: SpareParts

الجدول التالي يمثل بعض المقاييس التي تهمنا لمعرفة هل المتغيرات المستقلة لها تأثير كبير في النموذج أم لا هو:

جدول (10-7)

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.998[a]	.996	.995	1.36630

a. Predictors: (Constant), WorkHours

b. Dependent Variable: SpareParts

يظهر فيه معامل التحديد R وله قيمة عالية جدا ولدينا أيضا R Square المعدلة و قيمته 0.996 أي أن المتغير المستقل يقوم بتفسير ما يقارب 99% من النموذج. ولدينا خطأ التقدير وهو هنا خطأ صغير قيمته 1.36 وكلما اقتربت قيمته من الصفر كان أفضل.

جدول تحليل التباين:

جدول (10-8)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3385.066	1	3385.066	1813.333	.000[a]
	Residual	14.934	8	1.867		
	Total	3400.000	9			

a. Predictors: (Constant), WorkHours

b. Dependent Variable: SpareParts

من هذا الجدول نستطيع اختبار معنوية الانحدار وذلك باستخدام الاختبار التالي:

الفروض الإحصائية:

H_0: الانحدار غير معنوي.

H_A: الانحدار معنوي.

خطوات الاختبار:

نقارن P-value = .000 بمستوى المعنوية 0.05 نلاحظ أن القيمة P-value أقل من مستوى المعنوية.

القرار الإحصائي:

نقوم برفض الفرض العدمي وهذا معناه أن الانحدار معنوي (دال).

كما يمكننا معرفة مجموع المربعات التالية ودرجات حرياتها:

- مجموع مربعات البواقي SSE = 14.934 بدرجة حرية 8
- مجموع مربعات الانحدار SSR = 3358.066 بدرجة حرية 1
- مجموع المربعات الكلي SST = 4300.00 بدرجة حرية 9

وأيضا هناك متوسطات المربعات التالية:

- MSE = 1.867 متوسط مربع الخطأ.
- MSR =3385.066 متوسط مربع الانحدار

جدول المعالم المقدرة:

جدول (10-9)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	% Confidence Interval f B	
		B	Std. Error	Beta			ower Bound	pper Bound
1	(Constant	-4.758	1.357		-3.508	.008	-7.887	-1.630
	WorkHou	.498	.012	.998	42.583	.000	.471	.525

a. Dependent Variable: SpareParts

من هذا الجدول نستطيع الحصول على تقديرات لقيم معالم الانحدار المجهولة وبذلك نحصل على معادلة الانحدار المقدرة التالية:

$$\text{Spare Parts} = -4.758 + 0.498\text{Workhours}$$

حيث:

$$\hat{\beta}_0 = -4.758$$

$$\hat{\beta}_1 = 0.498$$

من معادلة الانحدار المقدرة بإمكاننا التنبؤ بقطع الغيار التي سنحصل عليها بناءا على ساعات العمل فمثلا إذا كانت ساعات العمل هي 70 ساعة فيكون الإنتاج المتوقع لقطع الغيار كالتالي:

$$\text{Spare Parts} = -4.758 + .498(70)$$
$$=30.102$$

أي تقريبا سيكون الإنتاج 30 قطعة غيار، كما يمكن من هذا الجدول معرفة الخطأ المعياري لكل من التقديرات و فترات الثقة لكل تقدير. و أيضا يمكن إجراء اختبارات المعنوية لكل من هذه التقديرات على حدة.

فلنختبر معنوية β_1 عند مستوى معنوية 0.05 :

الفروض الإحصائية:

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0$$

خطوات الاختبار:

نقارن P-value = .008 بمستوى المعنوية 0.05 نلاحظ أن القيمة P-value أقل من مستوى المعنوية.

القرار الإحصائي:

نقوم برفض الفرض العدمي وهذا معناه أن β_0 معنوي.

جدول البواقي:

جدول (10-10)

Residuals Statistics(a)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	20.1318	79.8682	50.0000	19.39377	10
Std. Predicted Value	-1.540	1.540	.000	1.000	10
Standard Error of Predicted Value	.433	.824	.597	.137	10
Adjusted Predicted Value	20.2070	79.7930	50.0073	19.35863	10
Residual	-2.44510	1.44949	.00000	1.28816	10
Std. Residual	-1.790	1.061	.000	.943	10
Stud. Residual	-1.936	1.143	-.003	1.026	10
Deleted Residual	-2.86228	1.68296	-.00726	1.52759	10
Stud. Deleted Residual	-2.485	1.169	-.067	1.153	10
Mahal. Distance	.003	2.372	.900	.852	10
Cook's Distance	.004	.320	.089	.103	10
Centered Leverage Value	.000	.264	.100	.095	10

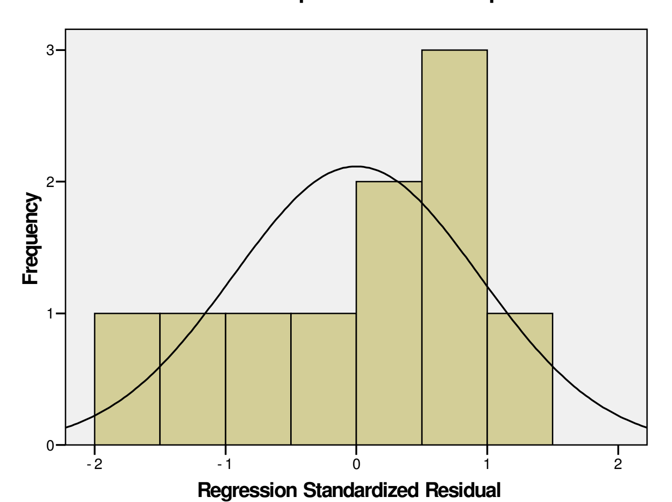
a Dependent Variable: Spare Parts

الأشكال البيانية:

- المدرج التكراري: ويستخدم لمعرفة إذا كانت البيانات تتوزع توزيعا طبيعيا أم لا، والمدرج التكراري الخاص بالنموذج هو:

Histogram

Dependent Variable: SpareParts



Mean =-2.22E-16
Std. Dev. =0.943
N =10

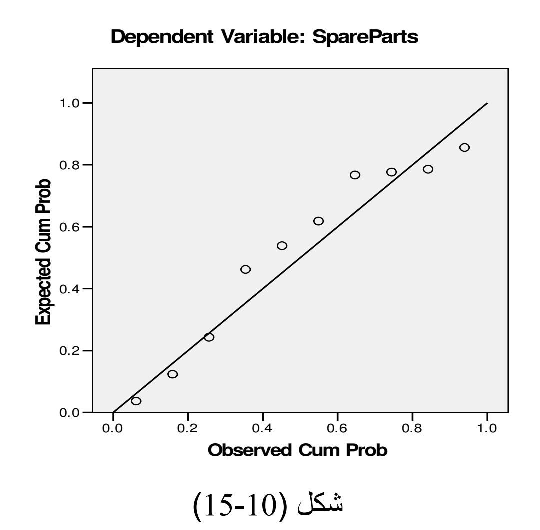
شكل (14-10)

نلاحظ أن شكله قريب من شكل التوزيع الطبيعي وهذا معناه أن البيانات تتبع التوزيع الطبيعي ونستطيع تطبيق تحليل الانحدار عليها.

رسم منحى الانحدار:

نشاهد أن النقاط تتوزع بشكل متساوي حيث إن النقاط الأعلى تساوي النقاط الأدنى من الخط ولذلك فإن فروض النموذج محققة فيه.

Normal P- P Plot of Regression Standardized Residual



شكل (10-15)

(3-10) الانحدار الخطي المتعدد Multiple Linear Regrassion

تحليل الانحدار المتعدد هو من بين كافة الأدوات الإحصائية، الأداة المستخدمة على أوسع نطاق، الفرق الأساسي بينه وبين الانحدار الخطي البسيط هو أنه هنا لدينا متغير تابع واحد و متغيرات مستقلة متعددة لها تأثيرات مختلفة على هذا المتغير التابع، ويكون شكل النموذج الخطي العام الذي يحتوي على عدد p-1 متغيرات مستقلة هو:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

حيث p هو عدد المعالم المجهولة في النموذج، ونلاحظ هنا أنه إذا وضعنا p=2 نحصل على النموذج الخطي البسيط.

تدعى المعالم $\beta_1, \beta_2, \dots, \beta_{p-1}$ بمعالم انحدار جزئية Partial regression Coefficients أو الميول الجزئية، لأن كل واحد منها يعكس تأثيرا جزئيا على المتغير التابع للمتغير المستقل الخاص به عندما يكون المتغير المستقل الآخر ثابتا.

إن فروض النموذج الخطي المتعدد هي نفس فروض النموذج البسيط يضاف إلى ذلك فرض عدم وجود ارتباط خطي متعدد بين المتغيرات المستقلة (Multicollinearity).

مثال (10-3)

نريد دراسة العلاقة بين ضغط دم الأطفال الرضع و عمرهم بالأيام و وزنهم عند الولادة مقاسا بالأونصة عن طريق نموذج انحدار خطي متعدد لستة عشر رضيعا باستخدام البيانات التالية:

جدول (10-11)

عدد	الوزن عند الولادة	العمر بالأيام	ضغط الدم
1	135	3	89
2	120	4	90
3	100	3	83
4	105	2	77
5	130	4	92
6	125	5	98
7	125	2	82
8	105	3	85
9	120	5	96
10	90	4	95
11	120	2	80
12	95	3	79
13	120	3	86
14	150	4	97
15	160	3	92
16	125	3	88

حيث نريد تقدير معاملات الانحدار لهذا النموذج أي إيجاد نموذج الانحدار المقدر عند مستوى معنوية 0.05.

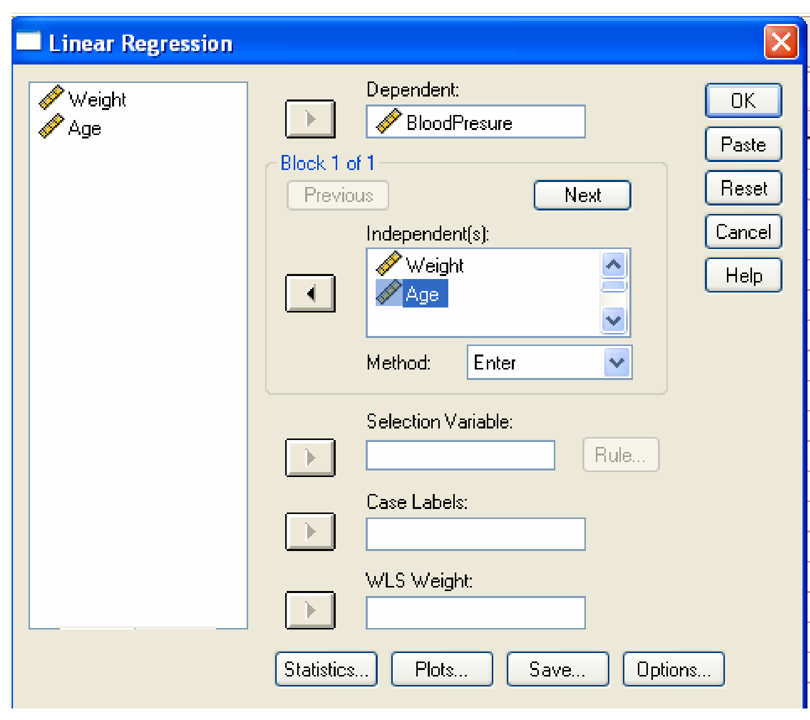
الحل:

نقوم بإدخال البيانات في ثلاثة أعمدة أحدها الوزن Weight، العمر Age وضغط الدم Blood Pressure كما يلي:

	Weight	Age	BloodPressure
1	135.00	3.00	89.00
2	120.00	4.00	90.00
3	100.00	3.00	83.00
4	105.00	2.00	77.00
5	130.00	4.00	92.00
6	125.00	5.00	98.00
7	125.00	2.00	82.00
8	105.00	3.00	85.00
9	120.00	5.00	96.00
10	90.00	4.00	95.00
11	120.00	2.00	80.00
12	95.00	3.00	79.00
13	120.00	3.00	86.00
14	150.00	4.00	97.00
15	160.00	3.00	92.00
16	125.00	3.00	88.00

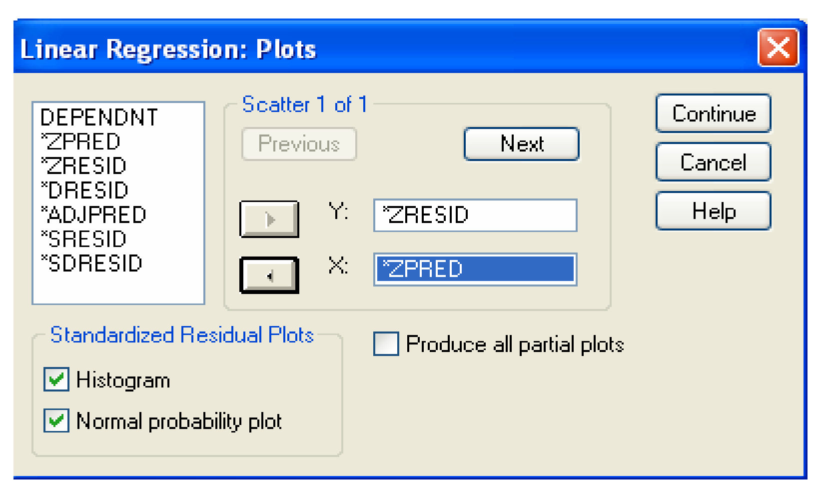
الشكل (10-16)

نقوم بتفعيل نافذة الانحدار شكل (10-17) ونختار المتغيرات المستقلة هي Age، Weight أما المتغير التابع فهو Blood Pressure كما يلي:



الشكل (10-17)

بعـد ذلـك نقـوم باختيـار الأمـر الفرعـي Statistics و نتأكـد مـن اختيـار كـل مـن Estimates،Confidence intervals، Covariance matrix ، Model fit ، Casewise diagnostics كما في الانحدار الخطي البسيط، ثم نعود إلى مربع الحوار السـابق بـالنقر علـى Continue ونذهب إلى الأمر الفرعي Plots و نتأكد من اختيار Histogram وNormal probability plot وأيضا نضع *ZRESID في المستطيل Y ونضع *ZPRED في المستطيل X كما يلي:



الشكل (10-18)

ثم نعود إلى مربع الحوار السـابق بـالنقر علـى Continue ثم ننقر OK فتظهر لدينا النتائج التالية:يظهر لدينا جدول يبين الطريقة المستخدمة في الانحدار والمتغير التـابع وهو Blood Pressure والمتغيرات المستقلة الداخلة في نموذج الانحدار الخطي وفي حالتنا يوجد متغيران و هما Age، Weight كما يلي:

جدول (10-12)

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Age, Weight[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: BloodPresure

بعض المقاييس التي تم حسابها للنموذج :

جدول (10-13)

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.939[a]	.881	.863	2.47917

a. Predictors: (Constant), Age, Weight

b. Dependent Variable: BloodPresure

لدينا المقياس R وكما نلاحظ فقيمته عالية وأيضا لدينا المقياس R Square والذي قيمته تقريبا 88. وقيمته المعدلة 0.863 أي أن المتغيران المستقلان في هذا النموذج لهما تأثير بمقدار 86.3% على المتغير التابع كما أنه لدينا الخطأ المعياري للتقدير في النموذج وهو 2.48 تقريبا.

جدول تحليل التباين:

جدول (10-14)

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	591.036	2	295.518	48.081	.000[a]
	Residual	79.902	13	6.146		
	Total	670.938	15			

a. Predictors: (Constant), Age, Weight

b. Dependent Variable: BloodPresure

وهو يظهر لنا قيم مجاميع المربعات و درجات الحرية الخاصة بكل مجموع مربعات بالإضافة إلى متوسطي مربعات الخطأ والبواقي، كما نلاحظ أنه عند إجراء اختبار المعنوية عند مستوى 0.05 نجد أن نموذج الانحدار المتعدد معنوي من آخر قيمة في الجدول.

جدول المعالم المقدرة:

جدول (15-10)

Coefficients(a)

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B Lower Bound	95% Confidence Interval for B Upper Bound
1	(Constant)	53.450	4.532		11.794	.000	43.660	63.241
	Weight	.126	.034	.352	3.657	.003	.051	.200
	Age	5.888	.680	.833	8.656	.000	4.418	7.357

a. Dependent Variable: BloodPresure

يختلف هذا الجدول عن جدول المعالم المقدرة الخاص بالانحدار الخطي البسيط بوجود أكثر من معلمتين فهنا لدينا ثلاثة معالم لنموذج الانحدار المتعدد وتكون معادلة الانحدار المقدرة:

$$\text{Blood Pressure} = 53.45 + 0.126\text{Weight} + 5.888\text{Age}$$

حيث:

$$\hat{\beta}_0 = 53.45$$
$$\hat{\beta}_1 = 0.126$$
$$\hat{\beta}_2 = 5.888$$

يمكنا أيضا القيام باختبارات المعنوية لهذه المعالم كل على حده بمقارنة آخر قيمة في الجدول مع مستوى المعنوية فنجد أن جميع المعالم معنوية.

جدول الارتباطات و التغايرات:

Coefficient Correlations(a)

جدول (16-10)

Model			Age	Weight
1	Correlations	Age	1.000	-.107
		Weight	-.107	1.000
	Covariances	Age	.463	-.002
		Weight	-.002	.001

a Dependent Variable: BloodPressure

يظهر لنا هذا الجدول قيم الارتباطات و التغايرات بين المتغيرات Weight،Age .

جدول البواقي:

Residuals Statistics(a)

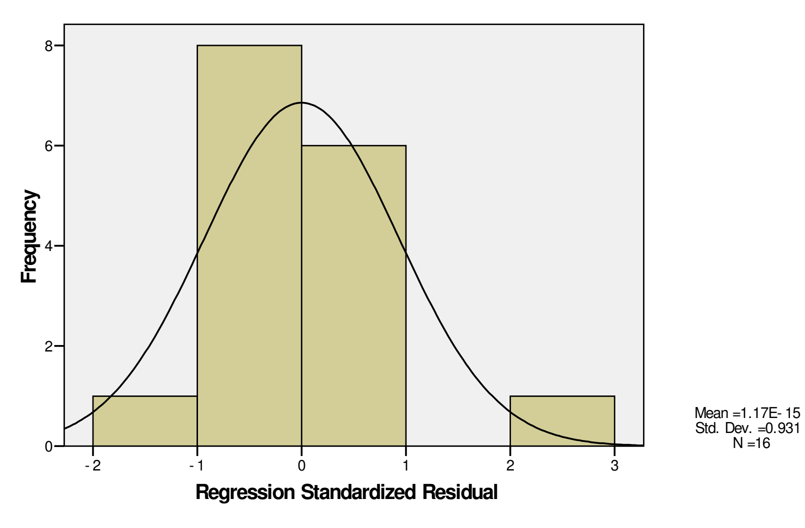
جدول (10-17)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	78.4119	98.5867	88.0625	6.27713	16
Std. Predicted Value	-1.537	1.677	.000	1.000	16
Standard Error of Predicted Value	.655	1.532	1.043	.261	16
Adjusted Predicted Value	78.8120	98.8089	87.9800	6.31927	16
Residual	-4.04376	6.69644	.00000	2.30798	16
Std. Residual	-1.631	2.701	.000	.931	16
Stud. Residual	-1.808	3.208	.014	1.074	16
Deleted Residual	-4.96979	9.44826	.08249	3.08420	16
Stud. Deleted Residual	-2.008	6.756	.223	1.871	16
Mahal. Distance	.109	4.793	1.875	1.358	16
Cook's Distance	.000	1.410	.123	.349	16
Centered Leverage Value	.007	.320	.125	.091	16

a Dependent Variable: BloodPressure

الأشكال البيانية:

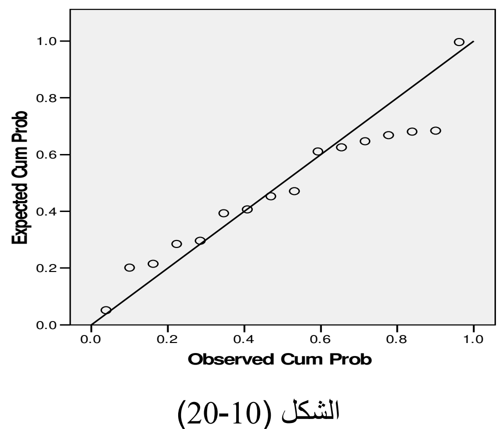
- رسم المدرج التكراري:



الشكل (10-19)

نستطيع منه القول بأن البيانات تتوزع وفق التوزيع الطبيعي.

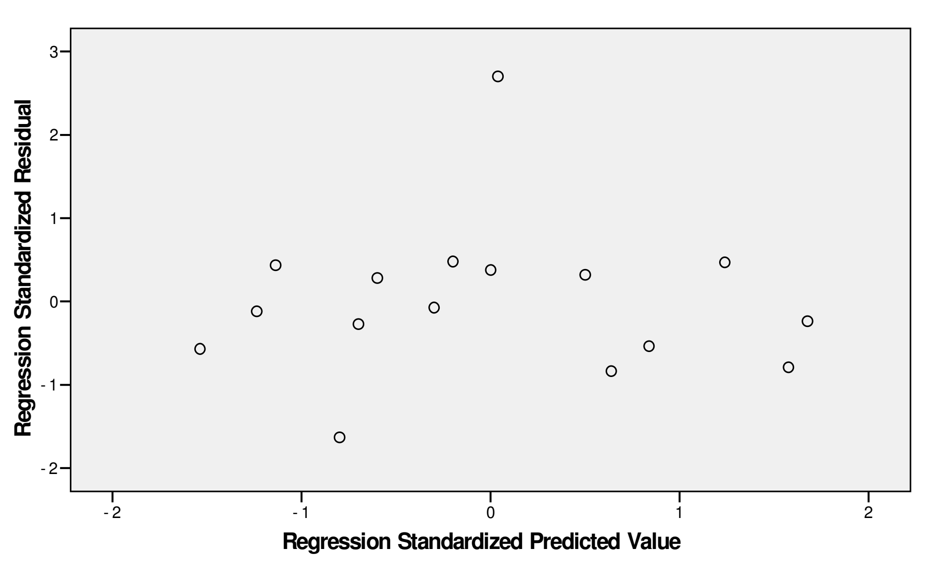
- رسم منحنى الانحدار:



الشكل (10-20)

هنا أيضا البيانات لها تقريبا نمط انحدار واضح على الرغم من وجود بعض القيم الشاذة عن خط الانحدار شذوذا كبيرا.

⊕ رسم انتشار البواقي:



الشكل (10-21)

من هذا الشكل نستطيع أن نقول أنه لا يوجد نمط معين لانتشار البواقي ضد القيم المتوقعة أي أن الانتشار عشوائي وهذا يتفق مع شرط الخطية المطلوب في الانحدار سواء كان بسيطا أم متعددا.

(4-10) اختيار المتغيرات المستقلة الداخلة في النموذج الخطي:

المعلومات النظرية التي ستستخدم لاختيار المتغيرات المستقلة قد يمكن تعيينها بسهولة في بعض الميادين، وغالبا ما يمكن القيام بتجارب تحت سيطرة المجرب تزوده ببيانات يمكن على أساسها تقدير معالم الانحدار واختبار الشكل النظري لدالة الانحدار. لكن في العديد من الميادين الأخرى مثل العلوم الاجتماعية، السلوكية، الصحية، والإدارية، يندر نسبيا وجود نماذج قابلة للاستخدام، ولمزيد من تعقيد الأمور قد تنطوي

النماذج النظرية المتوافرة على متغيرات مستقلة غير قابلة للقياس مباشرة مثل الدخول المستقبلية للأسرة في السنوات العشر القادمة، وتحت هذه الظروف يضطر الباحثون إلى توقع متغيرات مستقلة يتصورون أنها يمكن أن تكون على صلة بالمتغير التابع المراد دراسته، ومن الواضح أن هذه المتغيرات المستقلة قد تكون مجموعة كبيرة وليس بالضرورة أن يكون لها تأثير على المتغير المراد دراسته فيلجأ الباحثون إلى عدة طرق يمكن أن يقوموا من خلالها بتحديد المتغيرات التي لها تأثير على المتغير المدروس حتى يتمكنوا من التوصل إلى نموذج انحدار نهائي يحتوي أفضل المتغيرات التي لها تأثير في النموذج وسنقوم بدراسة أحد هذه الطرق.

- **الانحدار باستخدام الانحدار التدريجي.**

تقوم طريقة الانحدار التدريجي و التي تسمى أيضا انحدار الخطوة فخطوة إلى الأمام، بالإضافة إلى تقدير معالم نموذج الانحدار المقدر باختيار أفضل نماذج انحدار يمكن التوصل إليها من متغيرات مستقلة عدة مع متغير تابع، حيث أنه ليس بالضرورة أن تكون لجميع المتغيرات المستقلة التي اختارها الباحث نفس التأثير على المتغير التابع، وهذا يؤدي إلى أنه يمكن ترشيح عدة نماذج انحدار بمتغيرات مستقلة مختلفة ولنفس المتغير التابع على أنها تكون نماذج جيدة.

يقوم روتين الانحدار التدريجي أولا بتوفيق نموذج انحدار خطي بسيط لكل من المتغيرات X المرشحة وعددها p-1 حيث نقوم في كل نموذج انحدار من النماذج التي حصلنا عليها باختبار ما إذا كان الميل مساويا للصفر أم لا حيث نستخدم الاحصاءة التالية:

$$F^*_k = \frac{MSR(X_k)}{MSE(X_k)}, \quad k = 1,2,\dots,p-1$$

حيث أن $MSR(X_k)$ يقيس الانخفاض في التغير الكلي للمتغير التابع المصاحب لاستخدام المتغير X_k، والمتغير X_i الذي له أعلى قيمة في الاحصاءة يكون هو المتغير المرشح لأول إضافة في النموذج، حيث إذا تجاوزت قيمة F* مستوى محدد سلفا يضاف المتغير X، وفيما عدا ذلك ينتهي البرنامج معتبرا أنه لا يوجد أي متغير X مفيد بما يكفي لدخول نموذج الانحدار. لفرض أن المتغير X_7 هو المتغير الذي تم اختياره في أول محاولة. سيقوم البرنامج بعدها بتوفيق جميع نماذج الانحدار المتضمنة لمتغيرين

مستقلين أحدهما X_7، ولكل نموذج نقوم باختبار ما إذا كان الميل مساويا للصفر مستخدمين الاحصاءة:

$$F_k^* = \frac{MSR(X_k/X_7)}{MSE(X_k, X_7)}$$

فإذا تجاوزت قيمة F^* مستوى $F_{(1-\alpha,1,n-p)}$ الذي نحدده سلفا يضاف المتغير الثاني، وفيما عدا ذلك ينتهي البرنامج باختيار المتغير X_7 فقط.

لنفرض أن X_3 تمت إضافته في الخطوة السابقة عندئذ سيقوم روتين الانحدار التدريجي باختبار ما إذا كان ينبغي حذف أي من المتغيرات المستقلة الموجودة في النموذج، وبناءً عليه إما أن تحذف متغيرات أو يظل النموذج بالمتغيرات التي قد اختيرت سابقا، ثم نرجع مرة أخرى إلى إضافة متغيرات مستقلة في النموذج ثم إلى حذف المتغيرات كما في الخطوات السابقة إلى أن نصل إلى مرحلة لا يمكننا فيها إضافة أو حذف متغيرا مستقلة من النموذج وعندها تنتهي عملية البحث عن النموذج ويكون النموذج الموجود لدينا هو نموذج الانحدار المقدر بأفضل المتغيرات المستقلة.

مثال (10-4)

نريد معرفة المتغيرات التي لها تأثير على المصروفات المعيشية للأسرة قمنا باختيار 30 أسرة وسجلنا قيم المتغيرات التالية لها:

جدول (10-18)

X_{4i}	X_{3i}	X_{2i}	X_{1i}	Y_i	i
5	7	2	6	6.5	1
6	13	3	16	11.2	2
8	19	4	15	11.2	3
8	13	4	14	10.5	4
9	9.6	2	10	9.3	5
5	8	3	10	7.2	6
7	15	4	9	13.4	7
7	5	3	6	5	8
6	12	4	7	11.6	9
7	14	4	12	11.2	10
5	7	3	4	6.2	11
11	19	6	0	12	12

9	16	6	14	11.3	13
8	11	4	5	9.2	14
9	6	4	2	5.5	15
7	7.6	3	4	6	16
8	25	1	16	12.5	17
6	10	2	4	9.8	18
5	6.5	3	5	6.1	19
11	15.1	6	15	14.3	20
10	18	6	15	45.5	21
6	11	1	2	10.8	22
5	5.6	1	4	4.5	23
8	8.5	2	3	6.7	24
5	6.3	1	2	4.5	25
5	14	3	5	9.8	26
7	4.6	1	5	4	27
7	7.5	1	3	5.5	28
7	11.2	2	8	10	29
5	9.5	0	7	8.5	30

حيث:

Y_i : المصروفات المعيشية (ألف ريال سعودي).

X_{1i} : مستوى تعليم رب الأسرة (عدد سنوات دراسته).

X_{2i} : عدد الأطفال.

X_{3i} : دخل الأسرة (ألف ريال سعودي).

X_{4i} : عدد أفراد الأسرة.

نريد إيجاد نموذج الانحدار الخطي الخاص بالمصروفات المعيشية للأسر عند مستوى معنوية 0.05.

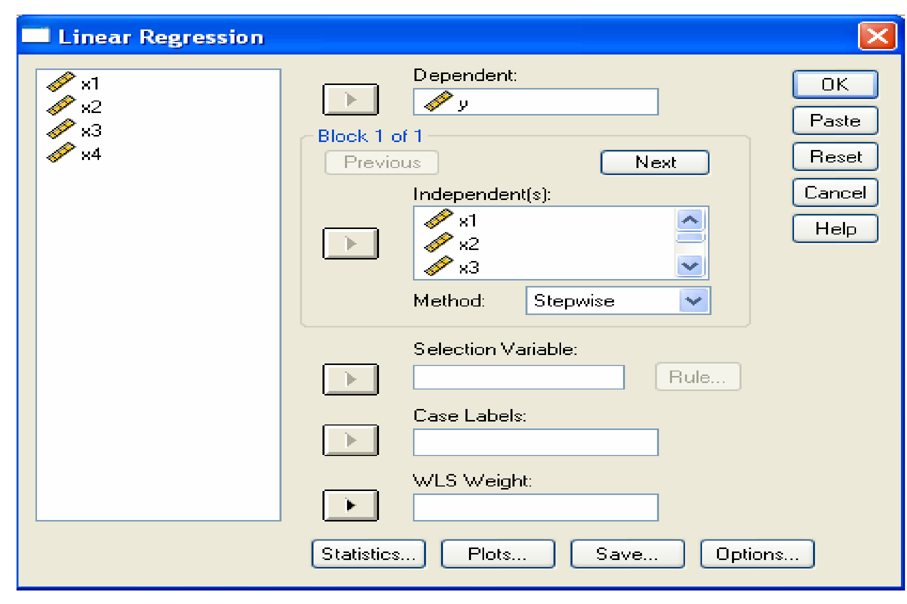
الحل:

ندخل البيانات في SPSSالبرنامج:

	y	x1	x2	x3	x4
1	6.50	6.00	2.00	7.00	5.00
2	11.20	16.00	3.00	13.00	6.00
3	11.20	15.00	4.00	19.00	8.00
4	10.50	14.00	4.00	13.00	8.00
5	9.30	10.00	2.00	9.60	9.00
6	7.20	10.00	3.00	8.00	5.00
7	13.40	9.00	4.00	15.00	7.00
8	5.00	6.00	3.00	5.00	7.00
9	11.60	7.00	4.00	12.00	6.00
10	11.20	12.00	4.00	14.00	7.00
11	6.20	4.00	3.00	7.00	5.00
12	12.00	.00	6.00	19.00	11.00
13	11.30	14.00	6.00	16.00	9.00
14	9.20	5.00	4.00	11.00	8.00

الشكل (22-10)

نقوم بتفعيل نافذة الانحدار ونختار المتغير التابع y والمتغيرات المستقلة هي x_1 x_2 x_3 و x_4 و نقوم باختيار الطريقة Stepwise كما يلي:



الشكل (23-10)

بالنسبة للخيارات Save، Plots،Statistics نقوم بما قمنا به في الانحدار المتعدد ثم ننقر ok فينتج لدينا:

جدول المتغيرات الداخلة في النموذج:

جدول (10-19)

Model	Variables Entered	Variables Removed	Method
1	x3	.	Stepwise (Criteria: Probability-of- Probability-of-،F-to-enter <= .050 F-to-remove >= .100).

Variables Entered/Removed(a)

a Dependent Variable: y

يتضح لنا من هذا الجدول أنه تم اختيار المتغير X_3 فقط في النموذج أي أن دخل الأسرة هو المتغير الوحيد المؤثر على المصروفات المعيشية من بين المغيرات التي قمنا باختيارها.

بعض مقاييس النموذج:

جدول (10-20)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.570(a)	.325	.301	6.11488

x_3،a Predictors: (Constant)

b Dependent Variable: y

نلاحظ من قيمة R Square المعدلة أن المتغير X_3 يفسر تقريبا 30% من Y.

جدول تحليل التباين:

ANOVA (b)

جدول (10-21)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	505.129	1	505.129	13.509	.001(a)
	Residual	1046.970	28	37.392		
	Total	1552.099	29			

x_3،a Predictors: (Constant)

b Dependent Variable: y

من هذا الجدول نستطيع اختبار معنوية الانحدار وذلك باستخدام الاختبار التالي:

الفروض الإحصائية:

H_0: الانحدار غير معنوي (غير دال).

H_A: الانحدار معنوي (دال).

جدول (10-2)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	.553	2.801		.197	.845	-5.184	6.290
	x3	.845	.230	.570	3.675	.001	.374	1.317

خطوات الاختبار:

نقارن P-value = .001 بمستوى المعنوية 0.05 نلاحظ أن القيمة P-value أقل من مستوى المعنوية.

القرار الإحصائي:

نقوم برفض الفرض العدمي وهذا معناه أن الانحدار معنوي.

جدول المعاملات:

Coefficients(a)

a Dependent Variable: y

من جدول المعاملات نجد أن معادلة الانحدار المقدرة هي:

$$Y=0.553+0.845X_3$$

حيث تكون قيمة $\hat{\beta}_0 = 0.553$, $\hat{\beta}_3 = 0.845$ فلنختبر معنوية β_3 عند مستوى معنوية 0.05 :

الفروض الإحصائية:

الفرض العدم هو: $H_0 : \beta_3 = 0$

ضد الفرض البديل: $H_A : \beta_3 \neq 0$

خطوات الاختبار:

نقارن *P-value = .001* مع مستوى المعنوية *0.05* نلاحظ أن القيمة *P-value* أقل من مستوى المعنوية.

القرار الإحصائي:

نقوم برفض الفرض العدمي وهذا معناه أن β_3 معنوي.

جدول المتغيرات المستبعدة من النموذج:

Excluded Variables(b)

جدول (21-10)

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	x1	.221(a)	1.120	.273	.211	.614
	x2	.312(a)	1.859	.074	.337	.786
	x4	.218(a)	1.211	.237	.227	.729

x3،a Predictors in the Model: (Constant)

b Dependent Variable: y

جدول البواقي:

Residuals Statistics(a)

جدول (22-10)

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4.4418	21.6882	9.9933	4.17352	30
Std. Predicted Value	-1.330	2.802	.000	1.000	30
Standard Error of Predicted Value	1.116	3.372	1.514	.457	30
Adjusted Predicted Value	4.4878	25.7033	10.0856	4.57232	30
Residual	-9.18822	29.72967	.00000	6.00853	30
Std. Residual	-1.503	4.862	.000	.983	30
Stud. Residual	-1.801	5.123	-.007	1.048	30
Deleted Residual	-13.20330	33.01107	-.09227	6.84940	30
Stud. Deleted Residual	-1.881	20.104	.491	3.727	30
Mahal. Distance	.000	7.852	.967	1.485	30
Cook's Distance	.000	1.448	.076	.290	30
Centered Leverage Value	.000	.271	.033	.051	30

a Dependent Variable: y

الفصل الحادي عشر

الاختبارات اللامعلمية (اللابارامترية)

Non-parametric Tests

(1-11) مقدمة:

في معظم الأساليب التي تكلمنا عنها في الاختبارات المعلمية (البلارامترية) نجد أنها مبنية على الفرضية التي تقول بأن العينة أو العينات العشوائية التي تم اختيارها للدراسة من مجتمع طبيعي، وغالبا ما تكون هذه الأساليب غير دقيقة إلى حد ما عندما يكون مجتمع العينة غير طبيعي، وحيث أن بعض المجتمعات لا تفي بالشروط المطلوبه لتطبيق تلك الأساليب، دعت الحاجة للبحث عن أساليب أخرى لا يتطلب تطبيقها مثل ذلك الشرط. هذه الأساليب يطلق عليها تسمية الأساليب اللامعلمية (اللابارامترية) لأنه وكما لوحظ في الفصل الأول كان اهتمامنا يركز على معلمة (بارامتر) أو أكثر من معلمات (بارامترات) المجتمع الإحصائي (المتوسط، التباين، النسبة،...الخ) علاوة على ذلك، وكما اشرنا في الفصل الأول، لكي نصل إلى استنتاج إحصائي يجب معرفة صيغة التوزيع الاحتمالي للمجتمع التي تم اختيار العينة منها.

وهناك نوعان من الأساليب الإحصائية تتم معاملتها على أنها أساليب لا معلمية وهما: أساليب لا معلمية بما تعنيه الكلمة وهي أساليب تختبر الفرضيات التي لا تتضمن أي نص يتعلق بمعلمات المجتمع الإحصائي، أما الأساليب الأخرى فهي أساليب التوزيعات الحرة وهي الأساليب التي لا تضع أي افتراضات على مجتمع العينة، و بصرف النظر عن التميز بين هذين الأسلوبين فإن كلاهما ستتم معاملتها على أنهما أساليب لا معلمية، هذه الأساليب يتم تطبيقها على سبيل المثال لا الحصر في الحالات التالية:

1-إذا كانت الفرضية المطلوب اختبارها لا تتضمن معلمة المجتمع.

2-البيانات مقاسه بمقياس أضعف من المقاييس المطلوبة لتطبيق الأساليب المعلمية مثل (المقياس الاسمي، المقياس الترتيبي، مقياس الفترة، المقياس النسبي).

3-إن لم تتوفر الشروط المطلوبة لتطبيق الأساليب المعلمية.

بعض مزايا الاختبارات اللامعلمية (اللابارامترية) ما يلي:

- مهما كان شكل التوزيع المأخوذ منه العينة فإن الاختبار اللامعلمي الذي له مستوى معنوية (دلالة) معين يكون له هذا المستوى فعلاً بشرط أن تكون العينة قد اختيرت عشوائياً كما يشترط أيضاً في بعض الحالات استمرار التوزيع.

- الإحصـاءات اللامعلميـة هـي الأسـلوب الوحيـد الممكـن اسـتخدامه فـي حالـة العينات الصغيرة جداً إلا إذا كان توزيع المجتمع معروفاً تماماً.
- يمكن استخدامها أحيانا للعينات التي تحتوي على مشاهدات من عدة مجتمعات متفاوتة.
- تصلح لتحليل البيانات التي تكون علي صورة رتب دون الحاجة إلى معرفة التوزيع في المجتمع الأصلي للبيانات.
- تستخدم في حالة كون البيانات تتضمن احدى صيغتي التفضيل مثلاً سليم أو معيب حيث السليم تكون له إشارة موجبة والمعيب إشارة سالبة وهنا لا تصلح الطرق التقليدية.
- يمكن تطبيقها عندما تكون البيانات مقاسه بمقياس ضعيف.
- تعتمد على افتراضات قليلة، وبالتالي فرصة تطبيقها خطأ ستكون صغيرة.
- الحسابات الضرورية للأساليب اللامعلمية عادة ما تكون سهلة ويمكن إنجازها بسرعة.
- سهولة فهمها وطريقة حسابها تجعلها مناسبة جدا للباحثين الذي ليس لهم خلفية علمية جيدة في الرياضيات و الإحصاء.

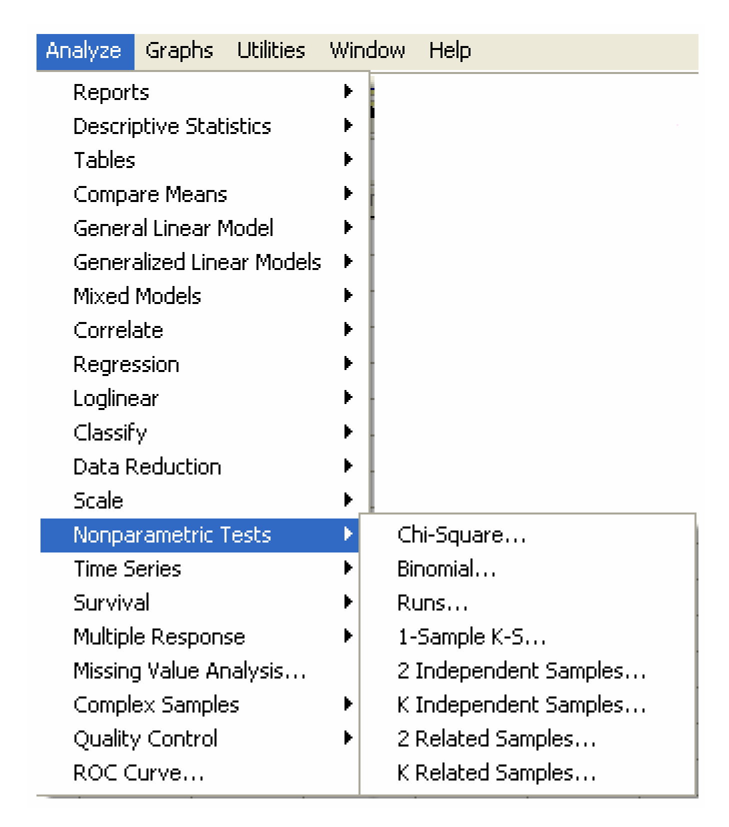
بعض عيوب الاختبارات اللامعلمية (اللابارامترية):

- نتيجة لسهولة حسابها، في بعض الأحيان يتم تطبيقها في مسائل يكون من الأفضل تطبيق أساليب معلمية عليها مما يسبب في ضياع المعلومات.
- في حالة العينات الكبيرة يؤدي استخدامها إلى جهد أكبر من الأساليب التقليدية.
- في حالة تحليل بيانات من توزيع طبيعي فإن استخدام الاختبارات اللامعلمية يعتبر فقد للبيانات وتقاس درجة الفقد بكفاءة الاختبار اللامعلمية.

وتنقسم الاختبارات اللامعلمية حسب عدد العينات عند إجراء الاختبار:

- حالة عينة واحدة One sample case
- حالة عينتين Two samples case (وهنا يوجد اختلاف بين المقياس التي تعتمد على: أ- العينتان مستقلتان ب- العينتان غير مستقلتين)
- حالة عدد العينات K والتي ربما تفترض استقلالا للعينات أو ارتباطها.

و تظهر أوامر الاختبارات اللامعلمية بالنقر على Analyze في شريط الأوامر فيظهر الشكل الآتي :



الشكل (1-11)

ومن الشكل السابق نستطيع اختيار نوع الاختبار الملائم لنوع المشكلة لدينا وسوف نتعرف فيما يلي على تلك الأنواع من الاختبارات وكيفية معالجتها للمشكلات المختلفة:

(2-11) اختبار مربع كاي (Chi-Square test):

إن من أشهر وأقدم اختبارات جودة المطابقة هو اختبار مربع كاي لجودة المطابقة، الذي اقترحه بيرسون (1900م) ،ويستخدم هذا الاختبار لتحديد ما إذا كانت التكرارات المشاهدة في جدول توزيع تكراري بسيط (لظاهرة واحدة) تتبع توزيع احتمالي معين مثل:

- اختبار أن عدد الحوادث التي تقع في ميدان معين لها توزيع بواسون (وهو أحد التوزيعات الاحتمالية المتقطعة شائعة الاستخدام في كثير من التطبيقات)،
- اختبار أن درجات الطلبة في أحد الامتحانات لها توزيع طبيعي، اختبار أن متوسط أطوال القطع التي تنتجها احدى الآلات لها توزيع طبيعي.

في كل هذة الحالات وأمثالها يقوم الاختبار أساسا على مقارنة التكرارات المشاهدة بالتكرارات المتوقعة التي تحسب باستخدام ذلك التوزيع الاحتمالي المعين (المطلوب اختبار ما إذا كانت البيانات تتبعه أم لا) ومن الفروق بين التكرارات المشاهدة والمتوقعة تحسب χ^2 بتطبيق الصيغة التالية:

$$\chi^2 = \sum \frac{(observed - \exp ected)^2}{\exp ected}$$

Observed: القيم المشاهدة

Expected: القيم المتوقعة

و إذا كان الجدول التكراري به m من الخلايا تكون χ^2 المحسوبة لها توزيع χ^2 بدرجات حرية يساوي عدد الخلايا مطروحا منها 1 أي m-1 ومستوى معنوية مقداره α وعليه إذا كان

$$\chi^2{}_{calcuate} > \chi^2{}_{(\alpha, m-1)}$$

يرفض الفرض العدم عند مستوى معنوية α حيث فرض العدم هنا هو فرض أن التكرارات المشاهدة بالجدول تتبع التوزيع الاحتمالي المعين.

هذا الاختبار يشبه اختبارات كاي للاستقلالية والتجانس من حيث كون إن إحصاءة الاختبار تنتج من مقارنة التكرارات المشاهدة ولكن أوجه تطبيقها مختلف تماما.

- **شروط تطبيق الاختبار:**

✓ تتضمن البيانات عينة عشوائية بها n من المفردات المستقلة عن بعضها البعض تم اختيارها من مجتمع x، ويمكن وضع هذه البيانات في جدول توافقي كما يلي:

جدول (1-11)

الصنف	$1\ \ 2\ \ 3......i\ ......\ r$	المجموع
التكرار المشاهد	$O_1\ \ O_2\ \ O_3.....O_i\ ......O_r$	n

حيث O_i تمثل عدد المفردات التي تقع في الصنف i حيث $i = 1,,2,3,...,r$ مع

ملاحظة أنه من الممكن أن يكون التصنيف نوعي أو كمي، فمثلا من الممكن تصنيف مجموعة من الأشخاص حسب الجنس (ذكور، إناث) أو ممكن التصنيف بالعمر...الخ.

- ✓ وحدة القياس على الأقل اسمية (nominal).

⌖ الفروض الإحصائية:

إذا رمزنا لدالة التوزيع غير المعروفة لمجتمع X بالرمز F(x) و لدالة التوزيع الفرضية بالرمز $F_o(x)$ وهي محددة بالكامل عدا من الممكن أن تكون المعلمة غير معروفة، ويجب تقديرها من بيانات العينة، فإنه يمكن صياغة الفرضيات الإحصائية كما يلي:

$H_0 : F(x) = F_0(x)$ لجميع قيم x

$H_A : F(x) \neq F_0(x)$ على الأقل لقيمة واحدة من قيم x.

⌖ إحصاءة الاختبار:

حيث أنه هناك احتمال بأن تقع أي مفردة يتم اختيارها من المجتمع بأي صنف من التصنيفات المختلفة و بالتالي يمكن الرمز لهذه الاحتمالات بالرمز $p_r....,\ p_2,\ p_1$.

على التوالي وذلك لأنه يوجد r صنفا، وعليه في حالة H_0 يمكن حساب التكرار المتوقع بكل صنف كما يلي:

$$T = \sum_{i=1}^{r} \frac{(O_i - E_i)^2}{E_i} \ \ \&\ \ E_i = np_i$$

حيث

O_i : القيم المشاهدة،

E_i: القيم المتوقعة،

P_i : قيمة الاحتمال،

n : مجموع المشاهدات.

- **القرار الإحصائي:**

نرفض H_o إذا كان $T > \chi^2_{\alpha,r-1}$

مثال (11-1)

أخذت عينة عشوائية مكونة من 400 أسرة من الأسر التي لكل منها ثلاثة أطفال أو أقل فوجد أن التوزيع التكراري لتلك الأسر حسب عدد الأطفال الذكور كالآتي:

جدول (11-2)

عدد الأطفال الذكور x	0	1	2	3
عدد الأسر	48	147	153	52

اختبر فرض أن عدد الأطفال الذكور بكل أسرة لديها ثلاثة أطفال له توزيع ذي الحدين بنجاح $\theta = 0.5$ وعدد المحاولات $N = 3$ وذلك عند مستوى معنوية مقدارها 0.05.

الحل:

- **أولا طريقة إدخال البيانات في البرنامج:**

نقوم بتصميم متغيرين أحدهم للمتغير الذي يصف عدد الأطفال في الأسرة و الآخر للمتغير الذي يصف التكرار المشاهد.

Number	observe
.00	48.0000
1.00	147.000
2.00	153.000
3.00	152.000

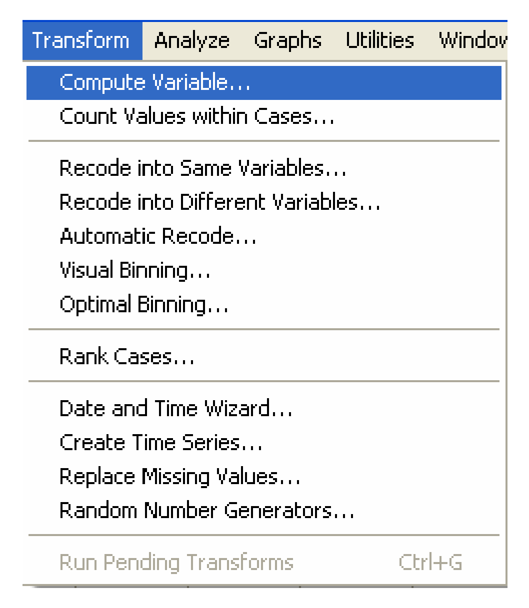
شكل (11-2)

- **ثانياً: نقوم بحساب التكرار المتوقع :**

ونلاحظ أننا نريد حساب التكرار المتوقع من توزيع ذي لحدين ولذلك يتم أولا حساب الاحتمالات والذي نرمز لها بالرمز p_i ومن ثم حساب التكرار المتوقع كالاتي:

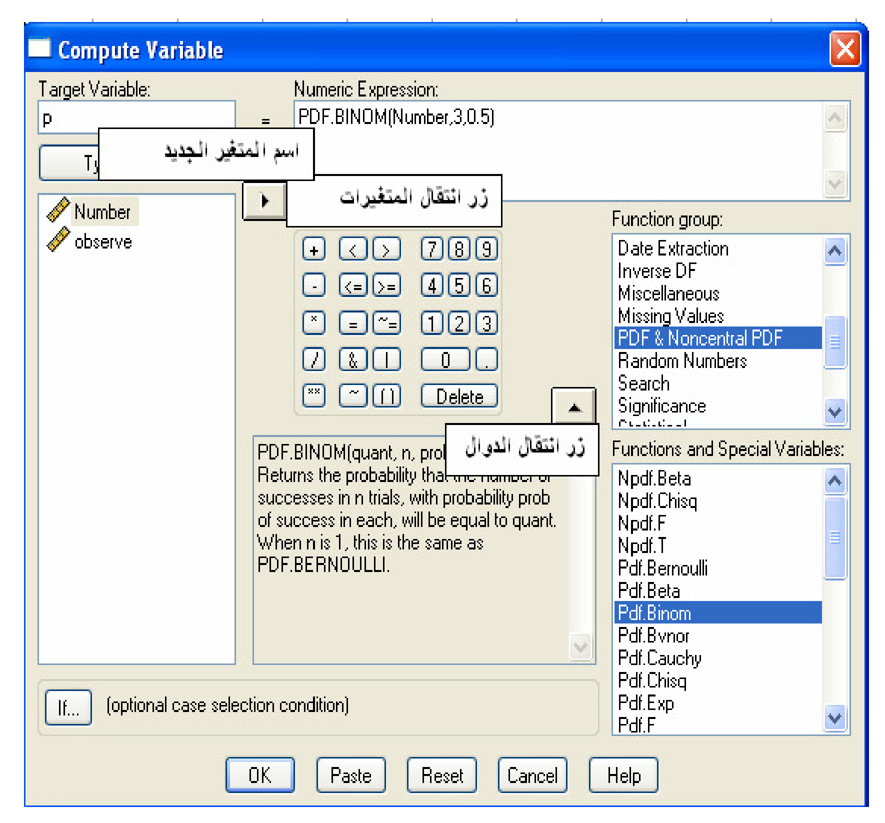
$$E_i = np_i$$
$$= 400 \ p_i$$

فمن قائمة Transform نقوم باختيار ....Compute Variable ومن ثم يظهر الشكل الآتي:



شكل (3-11)

ثم نقوم باختيار PDF and Noncentral PDF من عمود Function group وبعدها تظهر قائمة بدوال فرعية في عمود أسفل العمود السابق ونقوم باختيار pdf Binom وهذا يعني دالة الاحتمال لذي الحدين أي p_i فتضغط على الدالة مرتين متتالين أو النقر على زر الانتقال يظهر الشكل الآتي :



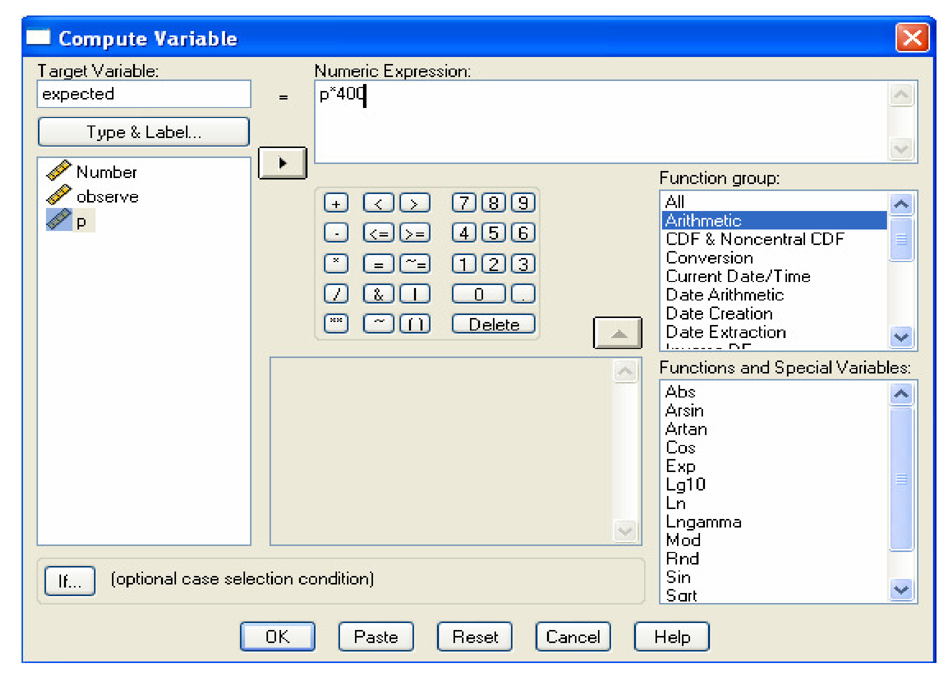
شكل (4-11)

ففي علامة الاستفهام الأولى نقوم بإدخال المتغير Number من زر إدخال المتغيرات بعد ذلك نقوم بإدخال عدد مرات المحاولة وهي في المثال تساوي 3 و أخيرا في علامة الاستفهام الثالثة و الأخيرة نقوم بإدخال احتمال الحدوث أو احتمال النجاح وهو في المثال 0.5. ونقوم تحت عنوان Target variable نقوم بإدخال اسم المتغير الجديد وهنا هو p الذي يرمز لدالة الاحتمال ومن ثم النقر على ok فتظهر شاشة Data View كالآتي:

جدول (3-11)

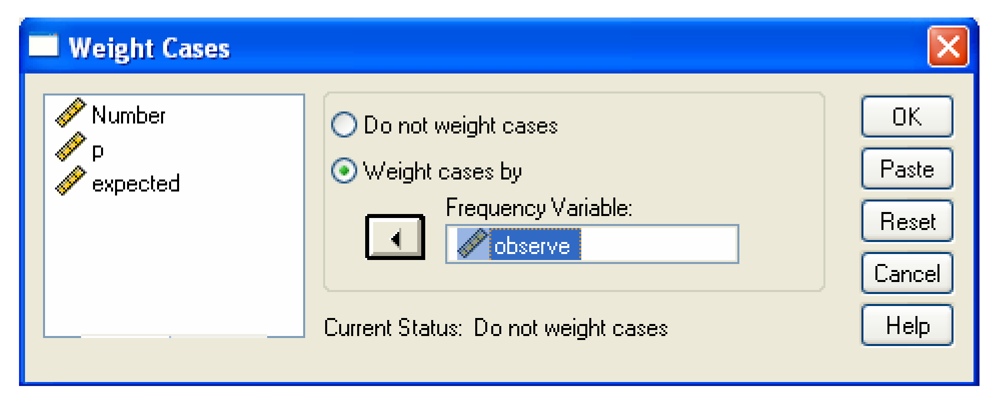
Number	observe	p
.00	48.0000	.13
1.00	147.000	.38
2.00	153.000	.38
3.00	152.000	.13

ومرة أخرى نقوم بحساب $np_i = 400 \ p_i$ بالطريقة السابقة ونسمي المتغير الجديد expected كما في الشكل الآتي:



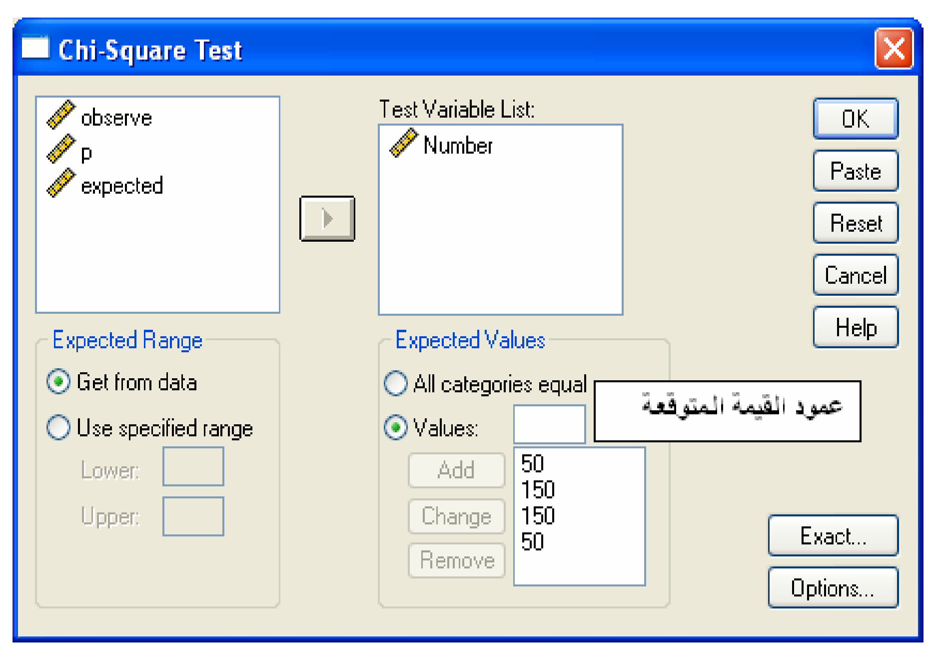
شكل (5-11)

بعد ذلك ننتقل إلى الخطوة الأخيرة حيث نقوم بتعريف أن العمود observed هو عمود تكرار العمود Number وذلك من خلال اختيار قائمة Data من شريط القوائم ومن ثم اختيار Weight cases ثم نقوم بقياس العمود الثاني كما في الشكل التالي:



شكل (6-11)

وبعد ذلك نقوم بفتح Analysis ثم Nonparametric ثم Chi square فيظهر الشكل الآتي:



شكل (7-11)

فنقوم بإدخال المتغير Number في عمود test variable list ونقوم بإدخال القيم المتوقعة المحسوبة في العمود الأخير في صفحة Data viewer على التوالي كما في الشكل الآتي حيث نقوم بتنشيط الخيار values من عمود Expected values و بعد ذلك نقوم بإدخال القيمة الأولى ثم ننقر على add ثم الثانية وهكذا و أخيرا نقوم بالنقر على ok في الشكل (7-11) وتخرج النتائج التالية:

جدول (4-11)

Number

	Observed N	Expected N	Residual
.00	48	50.0	-2.0
1.00	147	150.0	-3.0
2.00	153	150.0	3.0
3.00	52	50.0	2.0
Total	400		

جدول (11-4)

Test Statistics

	Number
Chi-Square[a]	.280
df	3
Asymp. Sig.	.964

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 50.0.

ويوضح الجدول (11-3) لمربع كاي القيمة المتوقعة لكل خلية.

القرار الإحصائي:

يوضح الجدول (11-4) قيمة الاختبار = 0.28 أيضا 964. = p-value وهي أكبر من مستوى المعنوية 0.05 إذا نقبل H_0 بأن عدد الأطفال الذكور بكل أسرة لديها ثلاثة أطفال له توزيع ذي الحدين بنجاح $\theta = 0.5$ وذلك عند مستوى معنوية مقدارها 0.05.

مثال (11-2)

اختيرت عينة عشوائية مكونة من 500 طالب من طلبة الفرقة الثالثة بإحدى كليات التجارة وتم تصنيفهم حسب التخصص ونتيجة الامتحان في الإحصاء فكان التصنيف كما يلي في الجدول الآتي:

جدول (11-5)

	ناجح P	راسب F	Total
إدارة الأعمال	240	60	300
اقتصاد	120	80	200
Total	360	140	500

اختبر ما إذا كان هناك علاقة بين التخصص ونتيجة الامتحان في الإحصاء عند مستوى معنوية 0.05.

الحل:

الفروض الإحصائية:

الفرض العدمي : لا توجد علاقة بين التخصص ونتيجة الإحصاء.

الفرض البديل : توجد علاقة بين التخصص ونتيجة الإحصاء .

⌖ أولا طريقة إدخال البيانات في البرنامج:

نقوم بإدخال رقم الصف الأول و العمود الأول ثم الصف الأول العمود الثاني ثم الصف الثاني العمود الأول ثم الصف الثاني العمود الثاني وبعد ذلك القيم تكون في عمود مستقل، ونعرف العمود الأول x حيث يشير إلى التخصص فيأخذ التخصص إدارة الأعمال (1) والتخصص الاقتصاد (2) و العمود الثاني y يشير إلى النتيجة في التخصص فتأخذ (1) في حالة النجاح وتأخذ (2) في حالة الرسوب انظر فصول أساسيات العرض و التحليل الإحصائي باستخدام البرنامج فيظهر الشكل التالي:

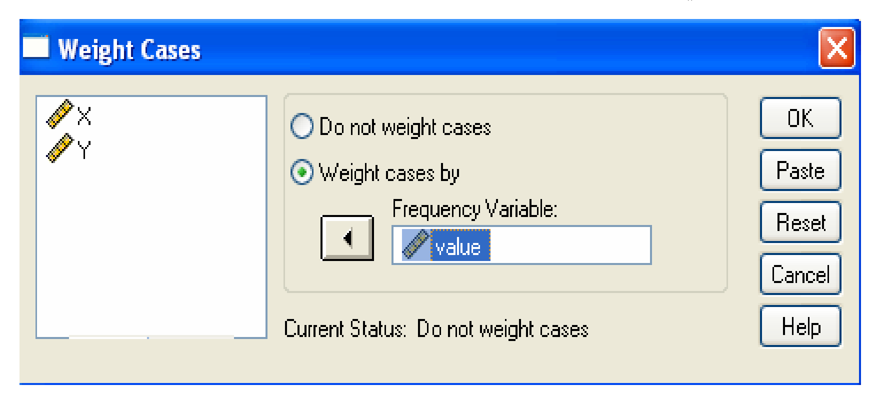
X	Y	value
1	1	240.00
1	2	60.00
2	1	120.00
2	2	80.00

شكل (11-8)

ثم نقوم بقياس القيم في العمود الثالث بالطريقة التالية:

Data→ Weight cases

فيظهر الشكل التالي:

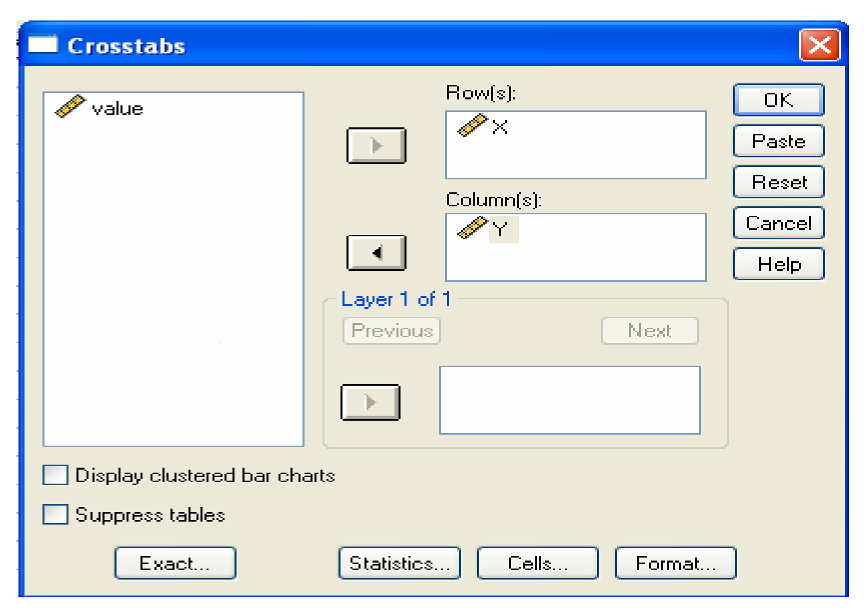


شكل (11-9)

ثم نقوم بالتالي:

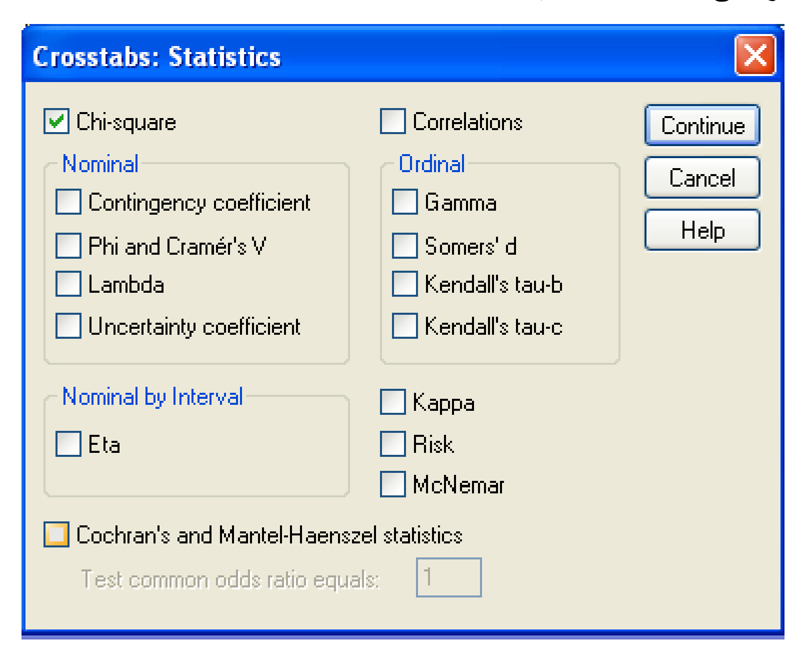
Analyze → Descriptive Statistics → Cross tabs ..

ثم ندخل العمود الأول في Rows و العمود الثاني في column كما يلي:



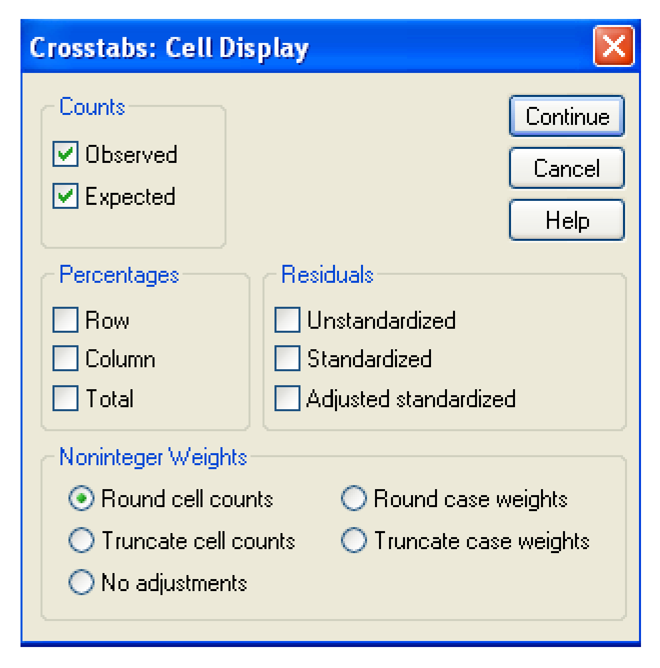
شكل (10-11)

ثم نقوم بالنقر على Statistics ونقوم باختيار Chi-square كما في الشكل التالي ومن ثم النقر على continue لتنشيط :



شكل (11-11)

ثم نقوم بالنقر على Cells ثم نقوم باختيار القيم المتوقعة لكل خلية كما في الشكل التالي:



شكل (11-12)

وبعد ذلك ننقر على Continue ثم OK في الشكل (11-11) و (11-12) فنحصل على النتائج التالية:

جدول (11-6)

X * Y Crosstabulation

			Y		Total
			1	2	
X	1	Count	240	60	300
		Expected Count	216.0	84.0	300.0
	2	Count	120	80	200
		Expected Count	144.0	56.0	200.0
Total		Count	360	140	500
		Expected Count	360.0	140.0	500.0

جدول (11-7)

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Squa	23.810[b]	1	.000		
Continuity Correcti[a]	22.828	1	.000		
Likelihood Ratio	23.507	1	.000		
Fisher's Exact Tes				.000	.000
Linear-by-Linear Association	23.762	1	.000		
N of Valid Cases	500				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected cou 00.

ويوضح جدول (11-6) القيمة المتوقعة لكل خلية.

القرار الإحصائي:

يوضح جدول (11-7) قيمة الاختبار = 23.810 و أيضـا p-value = .000 وهـي أقل من مستوى المعنوية 0.05إذا نرفض H_0 ومن ثم توجد علاقة بين التخصص ونتيجة الإحصاء .

مثال (11-3)

طرحت مائة فرشة أسنان جديدة على مائة رجل و مائة سيدة لاستخدامها ثم إبداء آرائهم هـل يفضـلون اسـتخدامها أم لا 32 مـن الرجـال و 26 مـن السـيدات أجـابوا بـأنهم لا يفضلون استخدام الفرشاة الجديدة. هل هذا يدل على وجود فرق في التفضيل بين الرجال و السيدات عند مستوى معنوية 0.05؟ حيث كانت البيانات كما بالجدول التالي:

جدول (11-8)

	يفضل	لا يفضل	Total
رجل	68	32	100
سيدة	74	26	100
Total	142	58	200

الحل:

الفروض الإحصائية:

فرض العدم : لا يوجد فرق في التفضيل بين الرجال و السيدات.

الرفض البديل : يوجد فرق في التفضيل بين الرجال و السيدات.

بنفس الطريقة السابقة نحصل على النتائج التالية:

جدول (11-9)

VAR00001 * VAR00002 Crosstabulation

			VAR00002		Total
			like	don't like	
VAR00001	man	Count	68	32	100
		Expected Count	71.0	29.0	100.0
	woman	Count	74	26	100
		Expected Count	71.0	29.0	100.0
Total		Count	142	58	200
		Expected Count	142.0	58.0	200.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.874[b]	1	.350		
Continuity Correction[a]	.607	1	.436		
Likelihood Ratio	.875	1	.349		
Fisher's Exact Test				.436	.218
Linear-by-Linear Association	.870	1	.351		
N of Valid Cases	200				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 29.00.

القرار الإحصائي:

في جدول (11-9) نلاحظ قيمة اختبار مربع كاي = 0.874 ونلاحظ أيضا قيمة p-value=0.35 وهي أكبر من 0.05 إذا نقبل H_0 ونقول أنه لا يوجد فرق في التفضيل بين الرجال و السيدات في تفضليهم للمنتج.

(3-11) اختبار الدورة (Run Test):

إن معظم الأساليب الإحصائية التي تستخدم لدراسة ظاهرة معينة بمجتمع ما، تفترض أن مفردات العينة التي يتم اختيارها من ذلك المجتمع ستكون عشوائية حتى يكون للاستنتاج الإحصائي له معنى فإذا وجد شك في عدم صحة هذا الافتراض يجب أن يكون لدينا أسلوب علمي للتحقق من ذلك، فمثلا عند استخدام أساليب مراقبة الجودة نقوم برسم خرائط للتحكم ودراسة عدد الوحدات المعيبة في الإنتاج، وللقيام بذلك عادة ما تؤخذ عينات من الإنتاج بشكل دوري ومعرفة عدد الوحدات المعيبة، وبالتالي قد يكون هذا العدد أكبر أو أصغر من العدد المسموح به وإن الهدف من وراء ذلك هو معرفة ما إذا كان هذا العدد من الوحدات المعيبة بالإنتاج التي ظهرت بالعينات المختارة يحدث بشكل عشوائي لأنه إذا لم يكن عشوائي فهو مؤشر على ضعف التحكم في الإنتاج، ولقد اقترح أسلوب لاختبار العشوائية في مثل هذه الحالة يطلق عليه اسم اختبار الدورة.

إن الأساليب أو الطرائق التي تستخدم لدراسة العشوائية بظاهرة معينة تعتمد أساسا على طبيعة وعدد الدورات الموجودة في بيانات تلك الظاهرة ، وتعرف الدورة على أنها متتابعة من العناصر المتشابهة تسبق وتلحق بعناصر من نوع آخر وعدد العناصر داخل كل الدورة يطلق عليه طول الدورة. وغالبا ما نشك في عشوائية السلسلة التي تمثل ظاهرة معينة وخاصة إذا كان عدد الدورات قليلا أو كثيرا.

- **شروط تطبيق الاختبار:**

تتضمن البيانات متتابعة من المفردات مرتبة على حسب حدوثها ، ويمكن تصنيفها إلى نوعين منفصلين فقط . ولنفرض إن n تمثل حجم العينة و n_1 تمثل عدد مفردات أو مشاهدات النوع الأول ، و n_2 تمثل عدد مفردات أو مشاهدات النوع الثاني.

- **الفروض الإحصائية:**

H_0: نتائج حدوث النوع الأول و الثاني من المفردات عشوائية.

H_A: نتائج حدوث النوعين غير عشوائية.

مثال (11-4)

نفرض أن لدينا البيانات التالية وهي خاصة بعدد الوحدات المعيبة خلال 20 يوم:

جدول (11-10)

الايام	عدد الوحدات	الايام	عدد الوحدات
1	8	11	15
2	5	12	12
3	9	13	9
4	14	14	8
5	9	15	7
6	3	16	6
7	5	17	5
8	10	18	3
9	9	19	2
10	3	20	1

و المطلوب معرفة هل هذه البيانات عشوائية أم لا وذلك عند مستوى معنوية (دلالة) 0.05؟

الحل:

الفروض الإحصائية:

الفرض العدمي : البيانات عشوائية.

الفرض البديل : البيانات غير عشوائية.

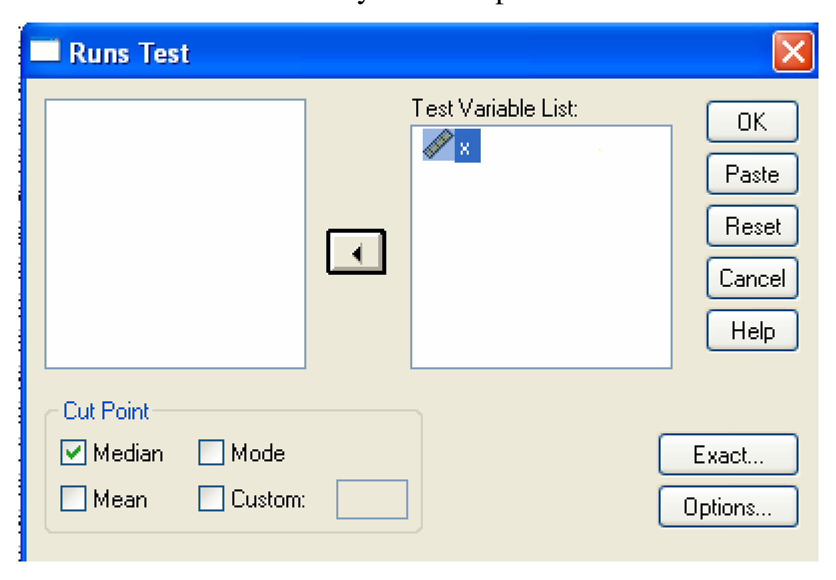
نقوم بإدخال البيانات في عمود متغير تم تسميته X كما هو موضح في الشكل التالي:

جدول (11-11)

	x
1	1.00
2	2.00
3	3.00
4	5.00
5	6.00
6	7.00
7	8.00
8	9.00
9	12.00
10	15.00
11	3.00
12	9.00
13	10.00
14	15.00
15	3.00
16	9.00
17	14.00
18	9.00
19	5.00
20	8.00

ثم نقوم بإجراء اختبار الدورة كما يلي:

Analyze→ Nonparametric Tests → Runs…



شكل (13-11)

ثم ننقر على OK فتكون نتيجة الاختبار كالتالي:

جدول (11-12)

Runs Test

	x
Test Value[a]	8.00
Cases < Test Value	9
Cases >= Test Value	11
Total Cases	20
Number of Runs	8
Z	-1.114
Asymp. Sig. (2-tailed)	.265

a. Median

من جدول المخرجات السابقة يتضح لنا أن قيم الوسيط والقيم الأكبر و الأصغر منها وعدد القيم الكلية و إحصائي الاختبار وهو Z.

القرار الإحصائي:

ونلاحظ أن p-value=.265 وهي أكبر من 0.05 يعني أننا نقبل الفرض الذي يقول إن البيانات عشوائية.

(4-11) اختبار كلوموجروف سيمنروف لعينه واحدة (1.Sample K-S):

اقترح العالم الروسي كلوموجروف في سنة 1933 اختبار جودة المطابقة في حالة عينة واحدة ، وفي سنة 1939 اقترح العالم الروسي سيمنروف اختبار جودة المطابقة في حالة بيانات تتعلق بعينتين، وبسبب وجود التشابه ما بين الاختبارين فقد أطلق على الاختبار الأول اسم كلوموجروف – سيمنروف لعينة واحدة وعلى الثاني اختبار كلوموجروف – سيمنروف لعينتين.

يعتمد هذا الاختبار على توزيعين احتماليين هما التوزيع الاحتمالي التراكمي النظري و التوزيع الاحتمالي التراكمي التجريبي ، فعند اختيار عينة عشوائية من مجتمع بتوزيع F(x) غير معروف حيث $F(x)=P(X\le x)$ فإن الهدف هو تحديد ما إذا كانت $F(x)=F_0(x)$ لجميع قيم x حيث $F_0(x)$ تمثل دالة التوزيع التراكمي الفرضية ، ولتحقيق هذا الهدف فإن اختبار كلوموجروف - سيمنروف لعينة واحدة ينظر إلى التقارب ما بين $F_0(x)$ و S(x) حيث S(x) تمثل دالة التوزيع التراكمي التجريبي ، فإذا كان هذا التقارب ضعيفا فإنه يعني عدم صحة الافتراض القائل بأن $F(x)=F_0(x)$ وخلاف ذلك الافتراض صحيحا.

يقوم هذا الاختبار بتحديد هل تتبع البيانات الداخلة توزيع معين أم لا، يستخدم الاختبار باختبار أربع أنواع من التوزيعات وهي:

- التوزيع الطبيعي.
- التوزيع المنتظم.
- التوزيع الآسي.
- توزيع بواسون.

- **شروط الاختبار**:

تتألف البيانات من عينة عشوائية $X_1,.....,X_n$ عدد مفرداتها يساوي n من مجتمع دالة توزيعه غير معروفة ونرمز لها بالرمز F(x).

⌖ **الفروض الإحصائية:**

$H_0 : F(x) = F_0(x)$ لجميع قيم x.

$H_1 : F(x) \neq F_0(x)$ على الأقل لقيمة واحدة من قيم x .

مثال (11-5)

البيانات التالية هي 30 مفردة لمتغير يمثل عدد ساعات تشغيل آلة معينة X تم إدخالها في الجدول لتالي:

جدول (11- 13)

المشاهدة	عدد الساعات	المشاهدة	عدد الساعات
1	7	16	1
2	4	17	2
3	4	18	1
4	4	19	2
5	3	20	4
6	22	21	22
7	4	22	9
8	0.4	23	6
9	9	24	8
10	0.5	25	6
11	5	26	4
12	7	27	7
13	2	28	2
14	3	29	0.47
15	8	30	7

المطلوب اختبار هل هذه البيانات تتبع التوزيع الأسي (وهو أحد التوزيعات الاحصائية المتصله الهامه) أم لا عند مستوى معنوية 0.05؟

الحل:

الفروض الإحصائية:

الفرض العدمي: البيانات تتبع التوزيع الأسي.

الفرض البديل: البيانات لا تتبع التوزيع الأسي.

نقوم بإدخال البيانات كالتالي:

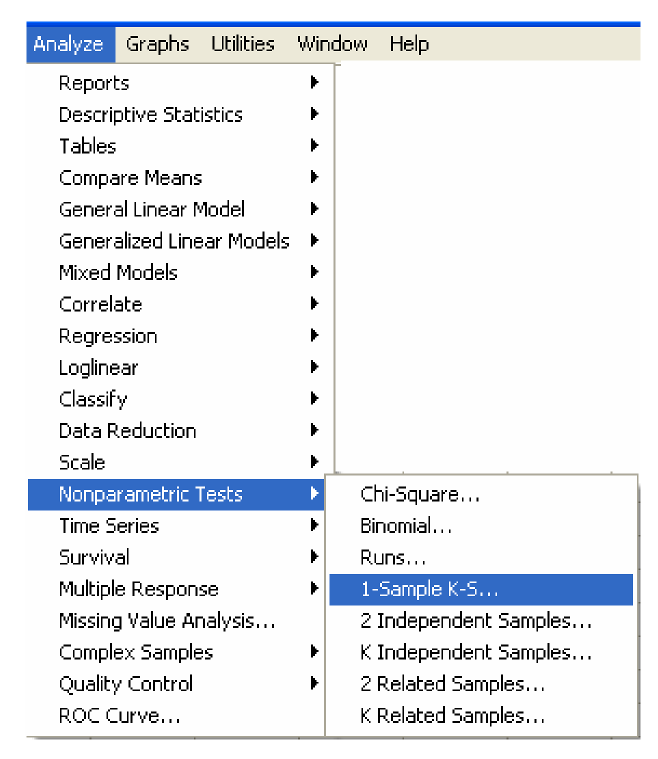
جدول (11- 14)

	X
1	7.00
2	4.00
3	3.00
4	4.00
5	9.00
6	5.00
7	2.00
8	8.00
9	2.00
10	2.00
11	22.00
12	6.00
13	6.00
14	7.00
15	.40
16	4.00
17	4.00
18	22.00
19	.40
20	.50
21	7.00
22	3.00
23	1.00
24	1.00
25	4.00
26	9.00
27	8.00
28	4.00
29	2.00
30	4.00

ثم نقوم بإجراء الاختبار كما يلي:

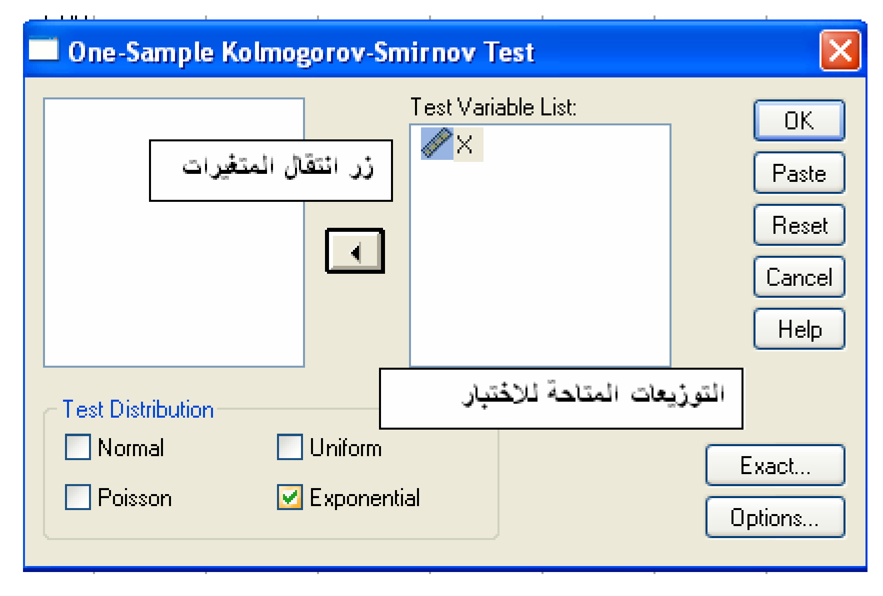
Analyze→ Nonparametric Tests → 1 sample K-S…

كما في الشكل الآتي:



شكل (11- 14)

فتظهر لدينا الشاشة التالية:



شكل (11- 15)

بعد النقر على OK تكون المخرجات كالتالي:

جدول (11- 15)

One-Sample Kolmogorov-Smirnov Test

		X
N		30
Exponential parameter[a,b]	Mean	5.3767
Most Extreme Differences	Absolute	.158
	Positive	.121
	Negative	-.158
Kolmogorov-Smirnov Z		.866
Asymp. Sig. (2-tailed)		.441

a. Test Distribution is Exponential.

b. Calculated from data.

المخرجات السابقة بها الوسط وبعض من المقاييس و قيمة دالة الاختبار و غيرها.

القرار الإحصائي:

نلاحظ أن p-value=.441 وهي أكبر من 0.05 إذا نقبل الفرض العدمي ونقول بأن البيانات تتبع التوزيع الأسي بمتوسط 5.3767 .

مثال (11-6)

البيانات التالية هي 20 مشاهدة لمتغير يمثل عدد المنتجات التي تم بيعها في احدى الشركات يوميا والتي نرمز لها بالرمز Y تم إدخالها في الجدول لتالي:

جدول (11- 16)

المشاهدة	عدد المنتجات	المشاهدة	عدد المنتجات
1	2	11	4
2	14	12	16
3	3	13	9
4	16	14	20
5	4	15	8
6	18	16	13
7	11	17	2
8	17	18	14
9	5	19	11
10	19	20	8

المطلوب اختبار هل هذه البيانات تتبع التوزيع الطبيعي أم لا عند مستوى معنوية (دلالة) 0.05؟

الحل:

الفروض الإحصائية

فرض العدم: البيانات تتبع التوزيع الطبيعي.

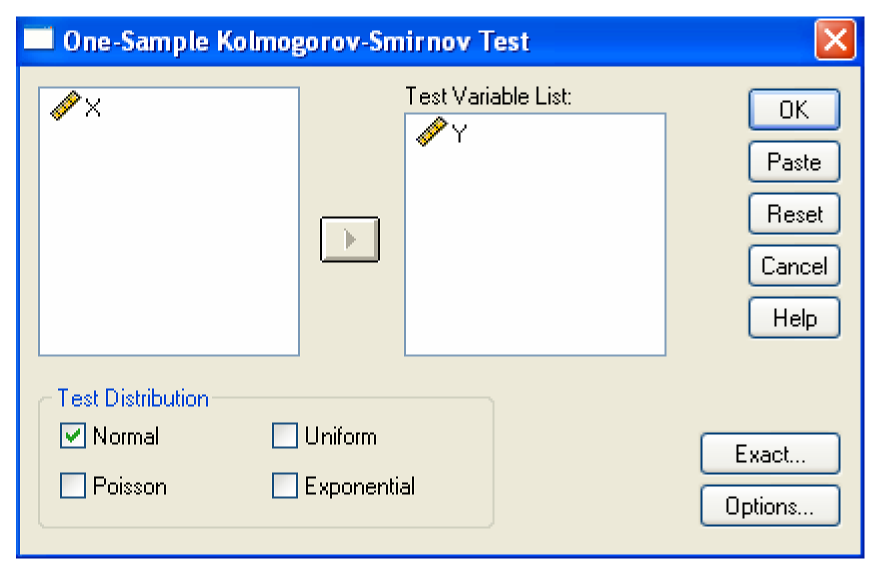
الفرض البديل: البيانات لا تتبع التوزيع الطبيعي.

نقوم بإدخال البيانات كالتالي:

جدول (11- 17)

Y
2.00
3.00
4.00
11.00
5.00
4.00
9.00
8.00
2.00
11.00
14.00
16.00
18.00
17.00
19.00
16.00
20.00
13.00
14.00
8.00

ثم نقوم بإجراء الاختبار باختيار التوزيع الطبيعي كالتالي:



شكل (11- 16)

بعد النقر على OK تكون المخرجات كالتالي:

جدول (11- 18)

One-Sample Kolmogorov-Smirnov Test

		Y
N		20
Normal Parameters[a,b]	Mean	10.7000
	Std. Deviation	5.99210
Most Extreme Differences	Absolute	.129
	Positive	.129
	Negative	-.112
Kolmogorov-Smirnov Z		.578
Asymp. Sig. (2-tailed)		.892

a. Test distribution is Normal.

b. Calculated from data.

المخرجات السابقة بها الوسط و الانحراف المعياري و غيرها.

القرار الإحصائي:

نلاحظ أن p-value=0.892 وهي أكبر من 0.05 إذا نقبل الفرض العدمي ونقول بأن البيانات تتبع التوزيع الطبيعي بمتوسط 10.7 وانحراف معياري 5.9921.

a) **اختبار مان – ويتني** (The Mann – Whitney test " U")

يستخدم هذا الاختبار لاختبار الفرضية H_0 التي تهدف إلى معرفة مدى تطابق مجتمعين من حيث معلمتي الموقع (المتوسط أو الوسيط) ، وذلك على أساس اختيار عينتين عشوائيتين منهما على أن تكون بيانات العينتين من نوع ترتيبي أي أنه يساعد في الإجابة على الأسئلة التي من النوع " هل أحد المجتمعين يبدو أنه يعطي قيم أكبر من المجتمع الآخر؟ " أو " هل وسيطي المجتمعين متساويين ؟". ويعتبر هذا الاختبار من أقوى الاختبارات اللامعلمية المستخدمة لهذا الغرض ، ويستخدم هذا الاختبار رتب المفردات نفسها ويفضل استخدام الرتب للأسباب التالية:

- إذا كانت الأعداد المعطاة للمفردات لا معنى لها بحد ذاتها ولكن يكون لها معنى في حالة مقارنتها بالترتيب مع الأعداد الأخرى فقط أي أن الأعداد لا تحتوي على معلومات أكثر مما تحتويه الرتب وهذا من طبيعة البيانات التي من نوع ترتيبي.
- حتى إذا كان لهذه الأعداد معنى ولكن دالة التوزيع لا تتبع التوزيع الطبيعي ، فان نظرية الاحتمالات عادة لا تكون في متناولها عندما تكون إحصاءة الاختبار تعتمد على البيانات الحقيقية، علاوة على ذلك فإن نظرية الاحتمالات المبنية على الرتب تعتبر نسبيا سهلة ولا تعتمد على التوزيع في كثير من الحالات.
- إن الكفاءة النسبية لاختبار مان-ويتني ليست بسيئة مقارنة باختبار t المألوف، وعليه يفضل استخدامه للأسباب المذكورة أعلاه.
- **شروط تطبيق الاختبار**:

1-تتضمن البيانات على عينة عشوائية من المفردات $X_1........,X_m$ من المجتمع "1" بدالة توزيع F وهذا يعني طبعا استقلالية المشاهدات وعلى عينة عشوائية أخرى $Y_1.......,Y_n$ من المجتمع "2" بدالة توزيع G و أيضا ذلك يعني استقلالية البيانات عن بعضها داخل العينة الواحدة.

2-العينتين مستقلتين عن بعضهما البعض.

3-وحده القياس على الأقل ترتيبي.

4-إذا وجد اختلاف بين دوال توزيع المجتمعين، فإن الاختلاف سيكون في موقع التوزيع، أي أنه إذا كانت $F(x) \neq G(x)$ (حيث F هي داله التوزيع الأول، G هي دالة التوزيع الثاني فإن

$$F(x) = G(x + \Delta)$$

أو

$$G(x) = F(x - \Delta)$$

وهذا يعني أن

$$Y \stackrel{d}{=} X + \Delta$$

حيث $\overset{d}{=}$ تشير إلى أنه من نفس التوزيع ونجد أيضا أن المعلمة Δ تشير إلى Location shift ومن ثم فان $\Delta = E(Y) - E(X)$.

⌖ **الفروض الإحصائية:**

$$H_0 : (E(X) = E(Y))$$

$$H_A : (E(X) \neq E(Y))$$

أو

$$H_0 = \Delta = 0$$

ضد الفرض البديل

$$H_A = \Delta \neq 0$$

⌖ إحصاءة الاختبار:

$$T = \sum_{i=1}^{m} R(X_i)$$

حيث:

$\sum_{i=1}^{m} R(X_i)$: مجموع الرتب لعينة المجتمع الأول.

m : حجم العينة الأولىو n: حجم العينة الثانية.

- القرار الإحصائي:

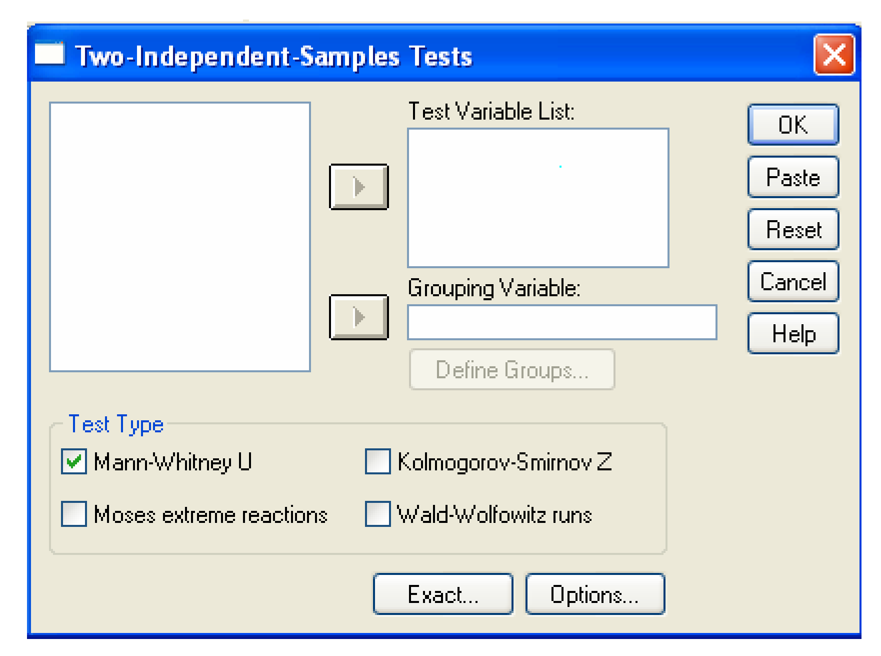
نرفض H_0 إذا كان

$$T \geq \omega_{\frac{\alpha}{2}} \quad or \quad T \leq n(m+n+1) - \omega_{\frac{\alpha}{2}}$$

ونلاحظ أن $\omega_{\frac{\alpha}{2}}$ يتم استخراجها من جدول التحليل اللامعلمي.

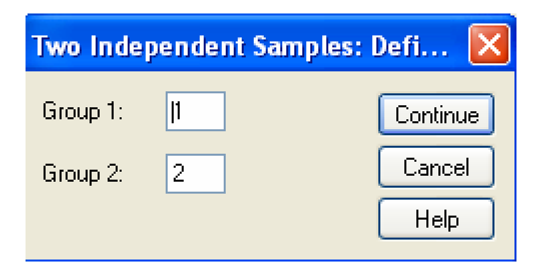
نقوم بتفعيل اختبار مان-ويتني كالأتي:

Analyze→ Nonparametric Tests →2 Independent Samples→ Mann – Whitney U



شكل (11- 17)

حيث يجب على الباحث أن يدخل قيم المتغير إلى test variable list وتصنيف المتغير في Grouping Variable حيث يعني أن "1" أن تلك القيمة تخص العينة الأولى، والقيمة "2" تعني أن القيمة تخص العينة الثانية كما في الشكل التالي:



شكل (11- 18)

مثال (11-7)

في دراسة لمقارنة بين متوسطي نسبة الملوحة في مياة نهر دجلة و مياة نهر النيل أخذت عينة على مدار 10 لمياة دجلة وعلى مدار 5 أيام لمياة النيل وسجلت النسب كما يلي:

مياة النيل العينة الأولى :

جدول (11- 19)

0.3	0.15	0.13	0.1	0.5

العينة الثانية عينة مياة نهر دجلة:

جدول (11- 20)

0.17	0.25	0.53	0.9	0.32	0.15	0.33	0.45	0.4	0.6

المطلوب اختبار أن نسبة الملوحة في المياة نهر دجلة هي نفسها نسبة الملوحة في نهر النيل عند مستوى معنوية 0.05 . مع العلم أن النسب مسحوبين من مجتمعين مستقلين؟

الحل:

الفروض الإحصائية:

H_0 : لا يوجد فرق في نسبة الملوحة بين المياة .

H_1 : يوجد فرق في نسبة الملوحة في النوعين.

نقوم بإدخال قيم المتغيرين في العمود الأول ونسميه X وقيم المتغيرين في العمود الثاني وفي هذا المثال يأخذ 1 و 2 و نسميه Group كما يلي:

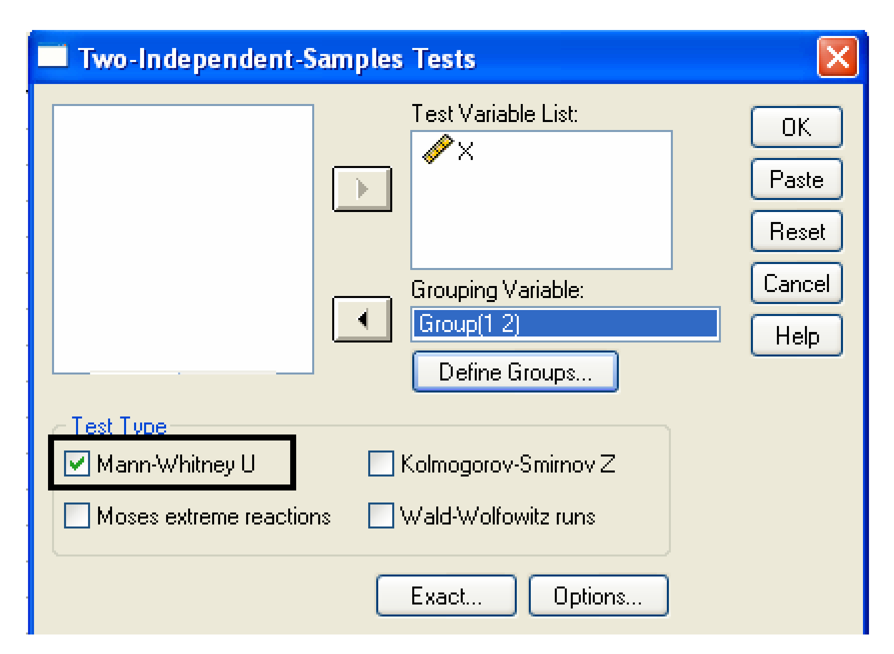
جدول (11- 21)

	X	Group
1	.50	1
2	.10	1
3	.13	1
4	.15	1
5	.30	1
6	.60	2
7	.40	2
8	.45	2
9	.33	2
10	.15	2
11	.32	2
12	.90	2
13	.53	2
14	.25	2
15	.17	2

ثم نقوم بإجراء اختبار عينتين مستقلتين مان ويتني:

Analyze→ Nonparametric Tests →2 Independent Samples→ Mann – Whitney U

ونضـع X فـي test variable list و Group فـي Grouping variable بعد تعريفها كما ذكرنا سابقا،كما في الشكل التالي:



شكل (11- 19)

بعد النقر على OK تخرج لنا النتائج التالية:

جدول (11- 23)

Ranks

	Group	N	Mean Rank	Sum of Ranks
X	1	5	5.10	25.50
	2	10	9.45	94.50
	Total	15		

جدول (11- 24)

Test Statistics [b]

	X
Mann-Whitney U	10.500
Wilcoxon W	25.500
Z	-1.777
Asymp. Sig. (2-tailed)	.075
Exact Sig. [2*(1-tailed Sig.)]	.075[a]

a. Not corrected for ties.

b. Grouping Variable: Group

الجدول الأول شكل (11-30) يحتوي على اسم المتغير وعن متوسط الرتب لكل متغير وغيرها.

⌖ **القرار الإحصائي:**

الجدول (11-24) يوضح بيانات إحصاءة الاختبار و هو اختبار مان ويتني ونلاحظ أيضا أن قيمة p-value = 0.075 وهي أكبر من $\frac{\alpha}{2} = .025$ إذا نقبل الفرض العدمي ونقول أنه لا يوجد فرق بين نسبة الملوحة بين النهرين.

مثال (11-8)

في تجربة لمقارنة لمتوسط الإنتاج اليومي من اللبن لنوعين من الأبقار أخذت عينة بحجم أربعة بقرات من النوع A وعينة بحجم خمس بقرات من النوع B وسجل الإنتاج اليومي كالتالي:

جدول (11- 25)

A	18	12	10	15	-
B	22	21	19	20	16

هل يمكن أن نستنتج أن الإنتاج اليومي لكل من النوعين متساوي عند مستوى معنوية 0.05؟

الحل:

الفروض الإحصائية:

H_0: لا يوجد اختلاف في الإنتاج اليومي لكل من النوعين.

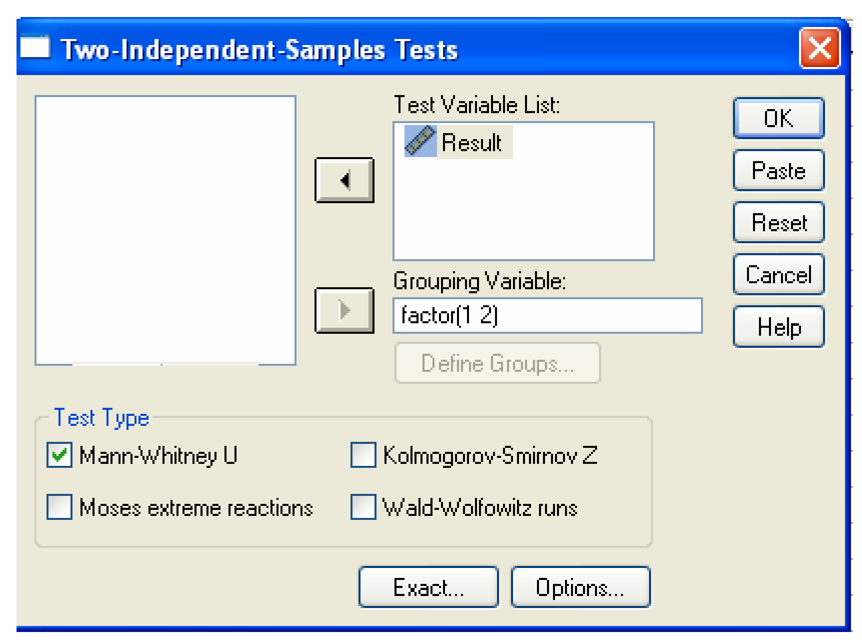
H_1 : يوجد اختلاف في الإنتاج اليومي لكل من النوعين.

نقوم بإدخال البيانات كما في السابق و كما هو موضح في الشكل التالي:

جدول (11- 26)

	factor	Result
1	1.00	18.00
2	1.00	12.00
3	1.00	10.00
4	1.00	15.00
5	2.00	22.00
6	2.00	21.00
7	2.00	19.00
8	2.00	20.00
9	2.00	16.00

ثم نقوم بإدخال كما في السابق وكما هو موضح في الشكل التالي:



شكل (11- 20)

بعد النقر على OK تخرج لنا النتائج التالية:

جدول (11- 27)

Ranks

	factor	N	Mean Rank	Sum of Ranks
Result	1.00	4	2.75	11.00
	2.00	5	6.80	34.00
	Total	9		

Test Statistics [b]

	Result
Mann-Whitney U	1.000
Wilcoxon W	11.000
Z	-2.205
Asymp. Sig. (2-tailed)	.027
Exact Sig. [2*(1-tailed Sig.)]	.032[a]

a. Not corrected for ties.

b. Grouping Variable: factor

من المخرجات نلاحظ مجموع ومتوسط الرتب لكل من المتغيرين وهذا في الجدول الأول (11-27).

القرار الإحصائي:

الجدول الثاني نلاحظ أن p-value= 0.027 وهي أكبر من $\frac{\alpha}{2} = .025$ (لأن الاختبار من طرفين) أي نقبل الفرض العدمي ونقول أنه لا يوجد اختلاف في الإنتاج اليومي لكل من النوعين.

b) **اختبار كرسكال والاس (The Kruskal – Wallis):**

يستخدم لاختبار الفرضية التي تهدف إلى معرفة ما إذا كانت عدة عينات قيد الدراسة قد تم اختيارها من مجتمعات بدوال توزيع متطابقة، علاوة على ذلك أنه يستخدم أكبر قدر ممكن من المعلومات التي بالعينات مقارنة باختبارات أخرى تستخدم لنفس الغرض، وعليه فهو أكثر قوة ويفضل استخدامه خاصة عندما تكون وحده قياس البيانات على الأقل ترتيبي.

- **شروط تطبيق الاختبار:**
  - تحتوي البيانات على مفردات k عينة عشوائية حجم كل عينة منهم $n_1......, n_k$ على التوالي.
  - المفردات مستقلة عن بعضها البعض خلال العينة ومن عينة إلى أخرى.
  - وحدة القياس على الأقل ترتيبي.
  - جميع دوال التوزيعات لجميع العينات $F_1, F_2, ..., F_k$ ترتبط بعلاقة الآتية:

$$F_j(t) = F(t - \tau_i), \qquad -\infty < t < \infty$$

حيث أن $j = 1, 2, ..., k$ والرمز τ يقرأ "تاو".

- **الفروض الإحصائية:**

H_0 : دوال التوزيع لجميع المجتمعات متطابقة او

$$H_0 = [\tau_1, ..., \tau_k \ \text{ all equal}]$$

H_A : ليس جميع دوال التوزيع متطابقة او

$$H_A = [\tau_1, ..., \tau_k \ \text{ not all equal}]$$

- **إحصاءة الاختبار:**

$$T = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

حيث:

$R_i = \sum R(X_{ij})$: مجموع الرتب المعطاة لمفردات العينة i

n : مجموع عدد المفردات في جميع العينات.

⌖ **القرار الإحصائي:**

إذا كانت حجم العينات صغيرة فإننا

نرفض H_0 إذا كان $T \geq \omega_{1-\alpha}$

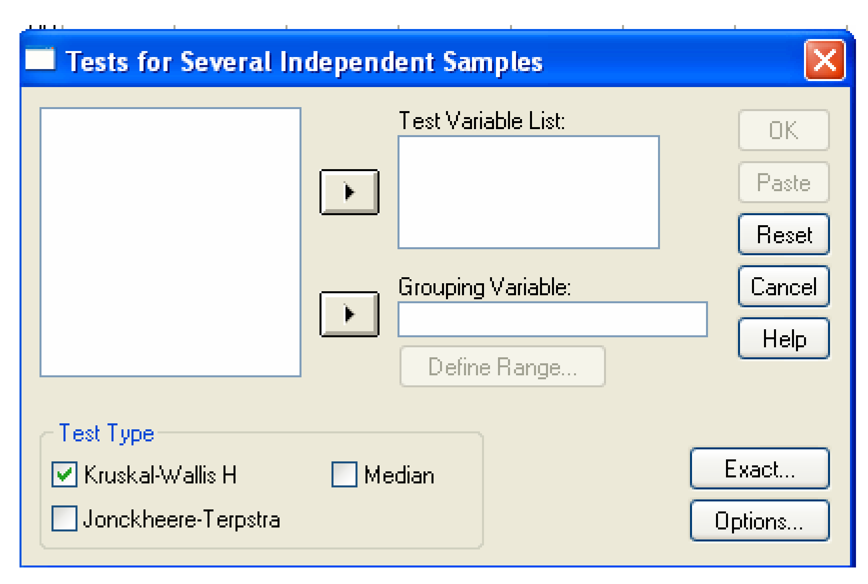
حيث تستخرج تلك القيم من جداول كرسكال والأس لتحليل اللامعلمي انظر Holiander and Wolfe (1998)

و إذا كانت حجم العينات كبيرة فإننا

نرفض H_0 إذا كان $T \geq \chi^2_{\alpha\,,\,k-1}$

نقوم بتفعيل اختبار كرسكال والأس كالآتي:

Analyze→ Nonparametric Test→ k independent Sample→ Kruskal - Wallis



شكل (11- 21)

حيث نقوم بإدخال عمود قيم المتغيرات في خانة تسمى Test Variable List و إدخال عمود تصنيف القيم "Group"في خانة Grouping Variable ثم نقوم بتعريفها بالنقر على مربع Define Range حيث يدخل الباحث رقم أول مجموعة دائما تكون واحد في خانة Minimum و رقم أخر مجموعة وهي عدد العينات في خانة Maximum كما في الشكل التالي:

Several Independent Samples: Define ...

Range for Grouping Variable

Minimum:

Maximum:

Continue

Cancel

Help

شكل (11 -21)

مثال (11-9)

تم اختبار ثلاثة أنواع من الأغذية على الأطفال ، تم اختبار الطعام A على عينة من 5 أطفال ، تم اختبار الطعام B على عينة من 6 أطفال ، تم اختبار الطعام C على عينة من 5 أطفال وتم قياس أوزانهم بعد 7 أسابيع وكانت النتائج كالتالي:

جدول (11 -28)

Food Type	أوزان الأطفال					
A	11.2	12.1	10.9	11.3	12	-
B	12.6	10.8	11.3	11	1.2	10.7
C	11.3	11.9	12.4	10.6	12	-

اختبر ما إذا كان متوسط أوزان الأطفال الذين تناولوا الأطعمة A،B ،C متساوية عند مستوى معنوية 0.05.

الحل:

- **الفروض الإحصائية:**

H_0 : جميع الأطعمة متساوية التأثير علي الوزن الأطفال.

H_A : على الأقل يوجد نوع من أنواع الأطعمة ذو تأثير متميز علي الوزن.

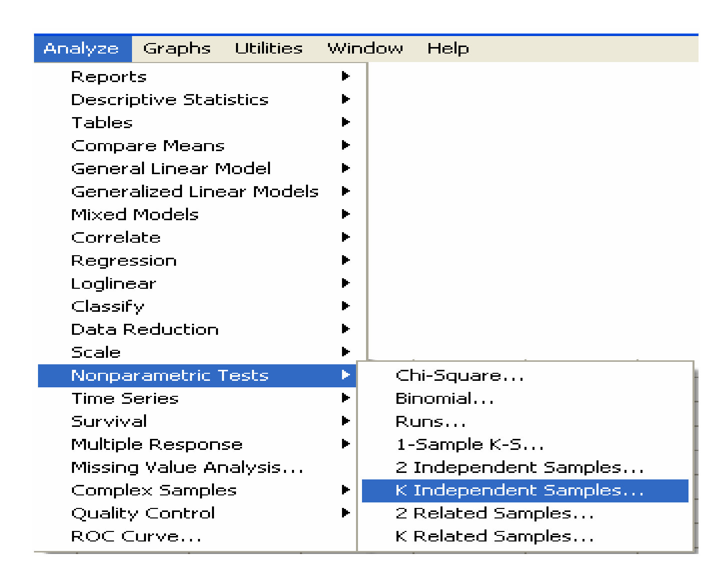
أولا نقوم بإدخال البيانات عمود يكون فيه قيم المتغيرات الثلاث و عمود يكون فيه تصنيف القيم كما يلي:

جدول (11- 29)

	Vaule	Group
1	11.30	1
2	11.90	1
3	12.40	1
4	10.60	1
5	12.00	1
6	12.60	2
7	10.80	2
8	11.30	2
9	11.00	2
10	12.00	2
11	10.70	2
12	11.20	3
13	12.10	3
14	10.90	3
15	11.30	3
16	12.00	3

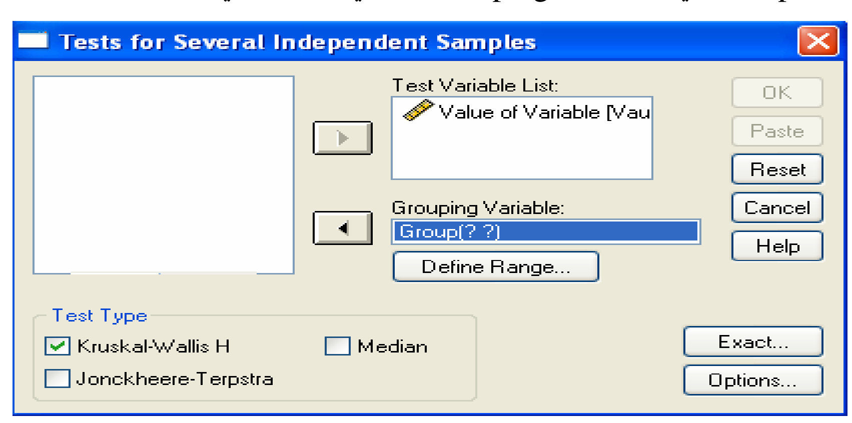
نقوم بالاتي:

- من شريط القوائم نختبار Analyze.
- ثم نختار Nonparametric Test.
- لاختبار اختبار أكثر من عينتين مستقلتين نختار k independent Sample كما في الشكل الآتي:



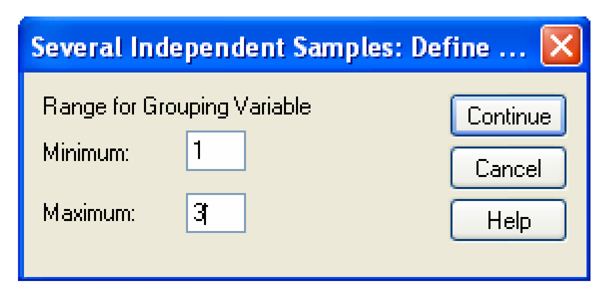
شكل (11- 22)

- فيظهر الشكل (11-20) ثم نقوم بختيار Kruskal – Wallis .
- نقوم بإدخال عمود Value في مربع Test Variable List
- و Group في Grouping Variable كما في الشكل التالي:



شكل (11-23)

- ونعرف المجموعة كما في الشكل التالي:



شكل (11-24)

⌖ ثم بعد النقر على OK تخرج لنا النتائج التالية:

جدول (11- 30)

Ranks

	Group	N	Mean Rank
Value of Variable	1	5	9.20
	2	6	7.67
	3	5	8.80
	Total	16	

جدول (11- 31)

Test Statistics[a,b]

	Value of Variable
Chi-Square	.315
df	2
Asymp. Sig.	.854

a. Kruskal Wallis Test

b. Grouping Variable: Group

نلاحظ من الجدول (11- 30) عدد العينة لكل مجموعة و أيضا متوسط الرتب لكل مجموعة.

والجدول (11- 31) هو أهم جدول لدينا لاحتوائه على معلومات الاختبار:

⌖ **القرار الإحصائي:**

نلاحظ من الجدول الثاني أن p-value= 0.854 وهو أكبر من مستوى المعنوية 0.05إذا نقبل الفرض العدمي ونقول أن تأثير الثلاثة أنواع من الأطعمة متساوية على الوزن.

مثال (11-10)

لمعرفة مدى تأثير ثلاث طرق إنتاج A، B، C لإنتاج منتج ما، تم عمل اختبار لتلك الطرق وكانت نتائج عدد الوحدات المنتجة في كل طريقة على مدار 5 أيام مختلفة كالتالي:

جدول (11- 32)

الطريقة A	13	15	25	15	17
الطريقة B	27	13	12	11	23
الطريقة C	14	16	18	19	20

اختبر صحة الفرض القائل بأنه على الأقل توجد طريقة من الطرق الثلاثة مؤثرة في حجم الإنتاج عند مستوى معنوية (دلالة) 0.01.

الحل:

⌖ **الفروض الإحصائية:**

H_0: الثلاثة طرق متساوية التأثير على الإنتاج

H_A: على الأقل يوجد طريقة واحدة لها تأثير مميز على حجم الإنتاج.

أولا نقوم بإدخال البيانات عمود يكون فيه قيم المتغيرات الثلاث و عمود يكون فيه تصنيف القيم كما في الشكل التالي:

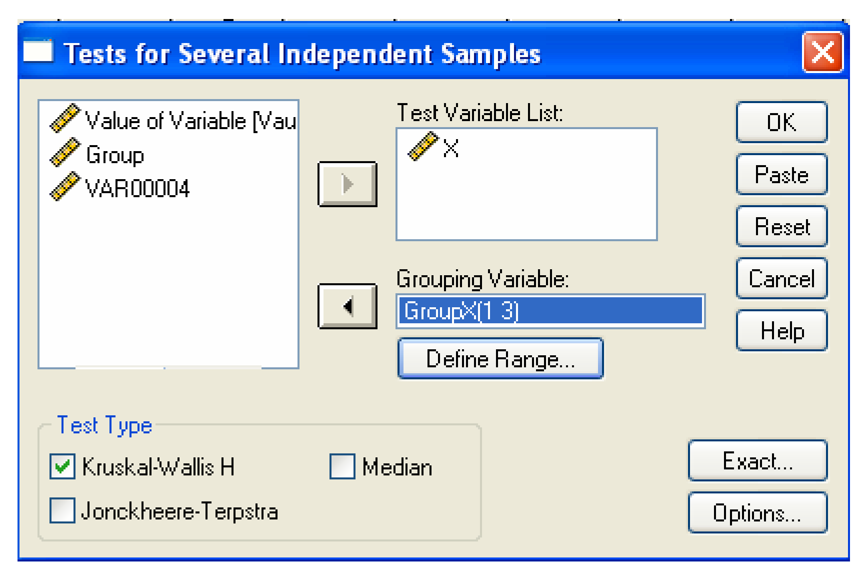
جدول (11- 33)

X	GroupX
13.00	1.00
15.00	1.00
25.00	1.00
15.00	1.00
17.00	1.00
27.00	2.00
13.00	2.00
12.00	2.00
11.00	2.00
23.00	2.00
14.00	3.00
16.00	3.00
18.00	3.00
19.00	3.00
20.00	3.00

نقوم بالاتي:

Analyze→ Nonparametric Test→ k independent Sample→ Kruskal - Wallis

ثم نقوم بإجراء اختبار عدة عينات مستقلة بإدخال X إلى test variable list و GroupX إلى Grouping variable كما في الشكل التالي:



شكل (11-25)

ثم بعد النقر على OK تخرج لنا النتائج التالية:

جدول (11- 34)

Ranks

	GroupX	N	Mean Rank
X	1.00	5	7.90
	2.00	5	6.90
	3.00	5	9.20
	Total	15	

جدول (11- 35)

Test Statistics[a,b]

	X
Chi-Square	.667
df	2
Asymp. Sig.	.716

a. Kruskal Wallis Test

b. Grouping Variable: GroupX

القرار الإحصائي:

نلاحظ من الجدول الثاني أن 716. =p-value وهو أكبر من مستوى المعنوية 0.01إذا نقبل الفرض العدمي ونقول أن الثلاثة طرق إنتاج لها تأثير متساوي على الإنتاج.

الفصل الثاني عشر

التعامل مع النتائج

Working with Output

(1-12) قص - نسخ -لصق جداول النتائج

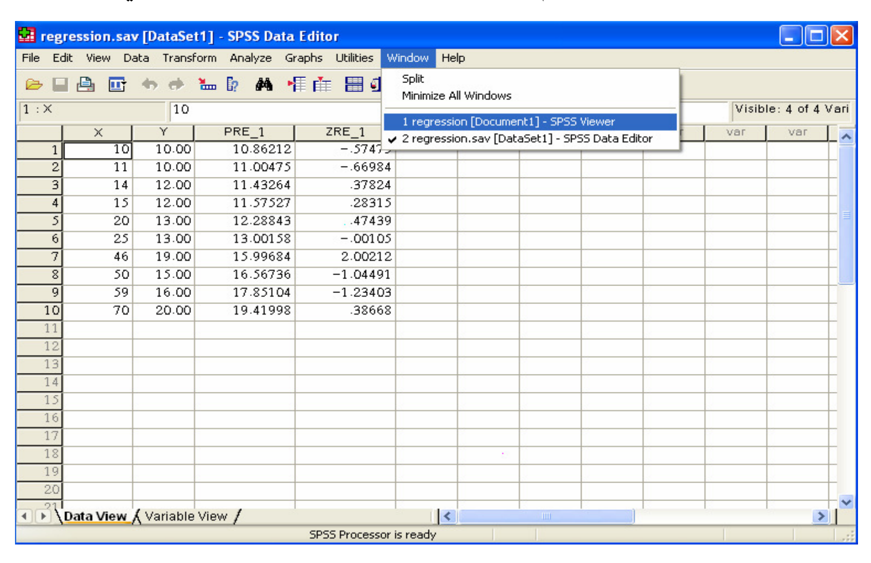
بعد إجراء التحليل الإحصائي سيكون لدينا نتائج ونريد نسخ تلك النتائج إلى ملفات تحرير النصوص مثل الملفات المصممة من خلال برنامج الورد. و هنا يتيح البرنامج لنا مجموعة من الخيارات مثل نسخ وقص ولصق الجدول والأشكال البيانية إلى ملفات تحرير النصوص لدينا المختلفة :

تطبيقات:

⌖ قص الجداول والأشكال البيانية:

مثال (1-12):

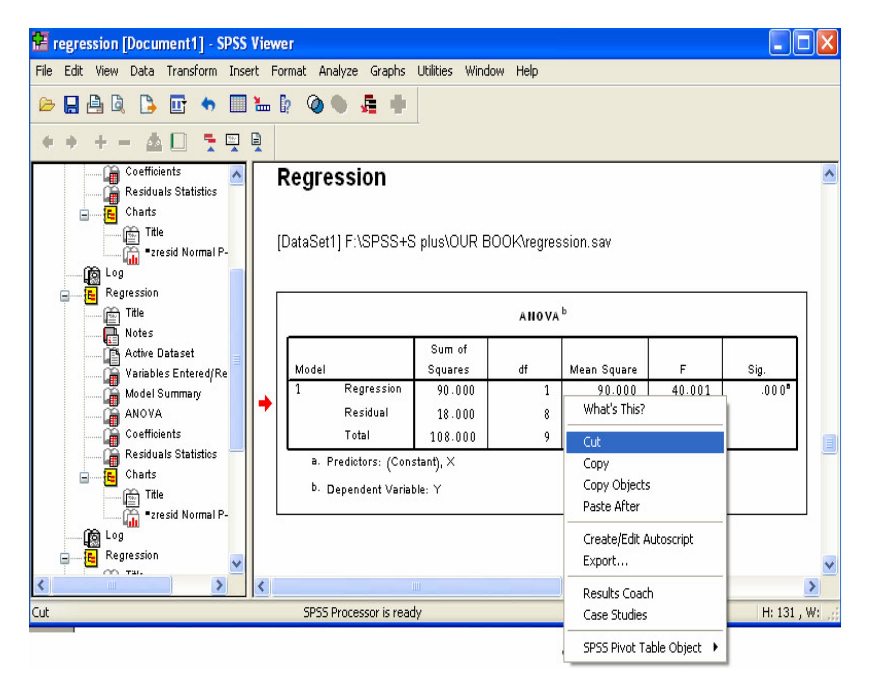
ولتحرير النتائج في مثال رقم (1-11) نقوم بفتح SPSS viewer وذلك من خلال النقر على Windows ثم اختيار SPSS viewer مثل الشكل الآتي:



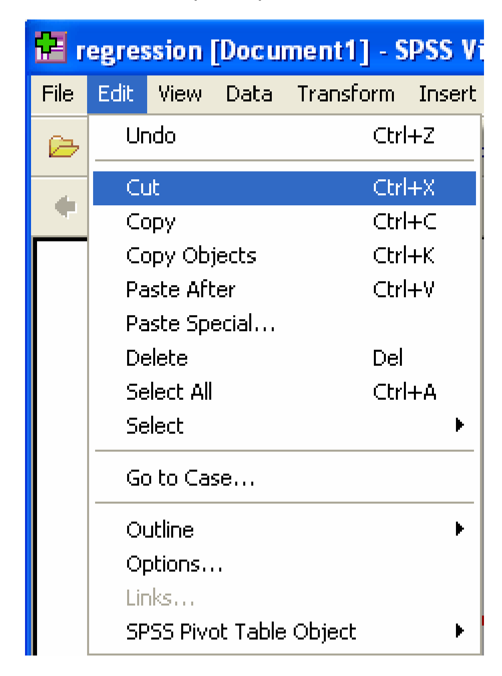
	X	Y	PRE_1	ZRE_1
1	10	10.00	10.86212	-.5747
2	11	10.00	11.00475	-.66984
3	14	12.00	11.43264	.37824
4	15	12.00	11.57527	.28315
5	20	13.00	12.28843	.47439
6	25	13.00	13.00158	-.00105
7	46	19.00	15.99684	2.00212
8	50	15.00	16.56736	-1.04491
9	59	16.00	17.85104	-1.23403
10	70	20.00	19.41998	.38668

شكل (1-12)

ومن ثم ظهور شكل محرر النتائج في الشكل (2-12) وإذا أردنا قص جدول تحليل التباين ANOVA فإننا ننقر بزر الماوس الأيسر عليها لتحديد الجدول وبعد ذلك ننقر بزر الفارة الأيمن ومن ثم الاختيار Cut أو بالنقر على edit من شريط القوائم ومن ثم اختيار cut كما في الشكل (3-12):



شكل (2-12)



شكل (3-12)

وبعد ذلك نذهب إلى ملف تحرير النصوص الذي نستخدمه ثم النقر على قائمة edit من شريط القوائم في برنامج التحرير ومن ثم اختيار past أو لصق على حسب نوع نسخة البرنامج عربي أو إنجليزي وكما فعلنا في الجداول يمكننا أيضا قص الأشكال البيانية بنفس الطريقة ويمكننا تلخيص الخطوات كالآتي:

- فتح صفحة النتائج أو محرر النتائج.
- في صفحة محرر النتائج انقر على الجدول المراد قصه.
- انقر على الزر الأيمن للفأرة في الجدول المراد تنسدل قائمة من الأوامر انقر على الأمر Cut.
- انتقل إلى الصفحة المراد لصق الجدول فيها وفي المكان المراد لصق الجدول فيه انقر على الزر الأيمن ثم انقر على الأمر Pasta .

(2-12) نسخ الجداول والأشكال البيانية:

هنا نقوم باتباع الخطوات الآتية:

- فتح صفحة محرر النتائج.
- في صفحة النتائج انقر على الجدول المراد نسخه.
- انقر على الزر الأيمن للفأرة تنسدل قائمة من الأوامر يظهر بها خيارين متعلقين ب النسخ Copy أو Copy object فإذا أردنا نسخ جدول بمحرره أي نسخه مع إمكانية التعديل في جدول النتائج نختار Copy وإذا أردنا نسخ الجدول كصورة نختار Copy objects والخيارات تلك ظاهرة في الشكل (2-12).
- انتقل إلى الصفحة المراد نسخ الجدول إليها وفي المكان المراد نقل الجدول إليه انقر على الزر الأيمن ثم انقر على الأمر Pasta .
- لاحظ أن الشكل المنقول غير منسق وغيابيا كان نوع الخط هو Arial من الحجم 10 يمكن باستخدام محرر Word تغيير نوع الخط مثلا / Time New Roman من الحجم 12 إذا تم اختيار الأمر Copy و إذا اخترنا Copy objects فلا يمكننا إجراء التعديل عليها فهي كصورة.

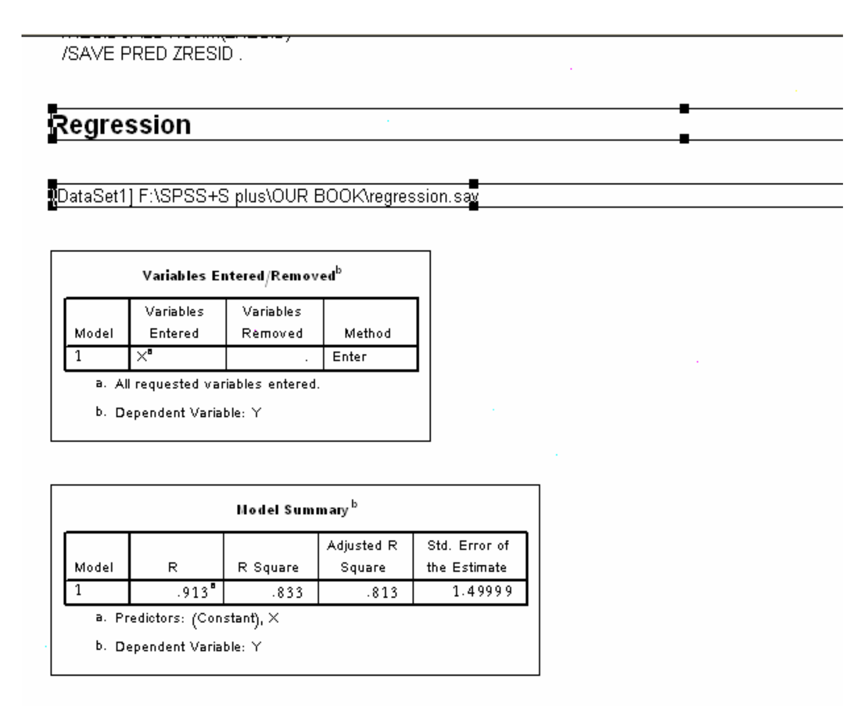
الآن نريد أن نتعامل مع المخرجات وهي الجداول بالإظهار والاختصار وإخفاء وإظهار بعض النتائج.

(3-12) الإظهار Expand والإخفاء Collapse وتحريك Move النتائج :

- **اظهار وإخفاء النتائج:**

إذا أردنا أن نخفي نتائج تمرين (1-11) نقوم بالآتي:

- انتقل إلى محرر النتائج
- نقوم تحديد صفحة النتائج من خلال النقر على edit ومن ثم اختيار Select all سوف تضاء كل محتويات الصفحة (يظهر إطار أسود لكل محتوى الصفحة انظر الشكل (12- 4):



/SAVE PRED ZRESID .

Regression

[DataSet1] F:\SPSS+S plus\OUR BOOK\regression.sav

Variables Entered/Removed [b]

Model	Variables Entered	Variables Removed	Method
1	X [a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

Model Summary [b]

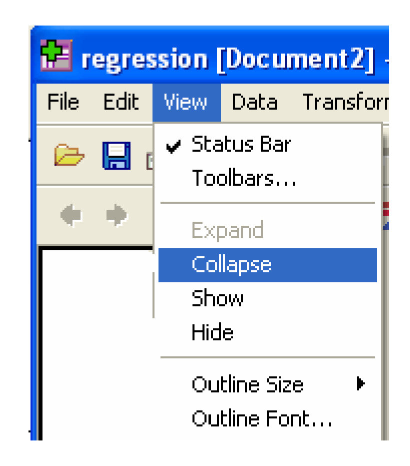
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.913 [a]	.833	.813	1.49999

a. Predictors: (Constant), X

b. Dependent Variable: Y

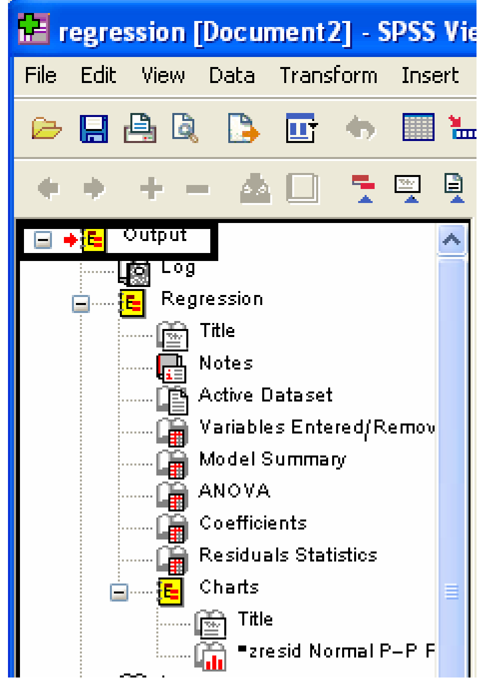
الشكل (12- 4)

- من القائمة View انقر الأمر Collapse ستختفي كل نتائج الصفحة عقب النقر انظر الآتي (انظر الشكل 12- 5):



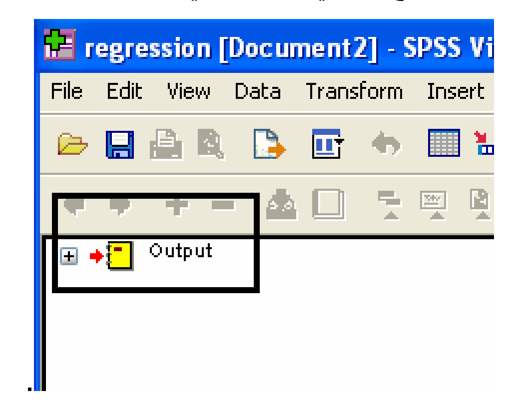
الشكل (12- 5)

- لإعادة اظهار المحتويات نكرر ما أجريناه في الخطوة السابقة مع النقر على الأمر Expand .
- ملاحظات:
- يمكن حذف محتويات الشاشة بالنقر على الأمر Delete من القائمة Edit ثم إعادتها مرة أخرى بالأمر Undo من نفس القائمة.
- يمكن الإخفاء بالنقر على المربع الموجود بداخله علامة الطرح في الجزء الأيسر يسار كلمة Output المشار إليه في الشكل الآتي:



الشكل (12- 6)

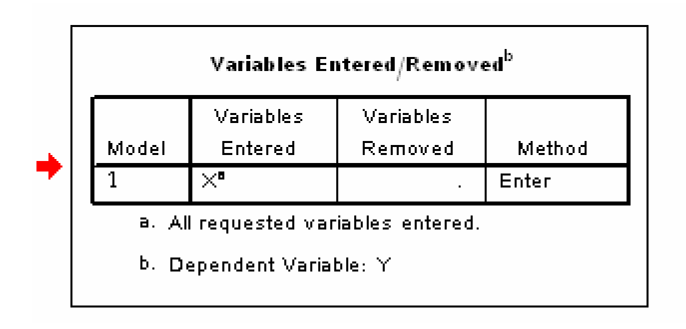
⌖ والإظهار بإعادة النقر مرة أخرى على نفس المربع ونلاحظ أنه علامة الطرح تحولت إلى علامة جمع كما في الشكل الآتي:



الشكل (12- 7)

- إخفاء جزء من عنوان Titled Sub-Item لنتائج موجودة في صفحة النتائج كالآتي:

⌖ انقر على الجدول أو الشكل المراد إخفاءه في محرر النتائج فمثلا في المثال (1-11) إذا أردنا إخفاء الجدول الذي بعنوان Variables Entered/Removed نقوم بتحديده وذلك من خلال النقر عليه فيظهر الشكل الآتي ونلاحظ أن أي جدول يتم تحديده يظهر محدد بإطار أسود وبجواره من ناحية اليسار سهم أحمر اللون:



Variables Entered/Removed[b]

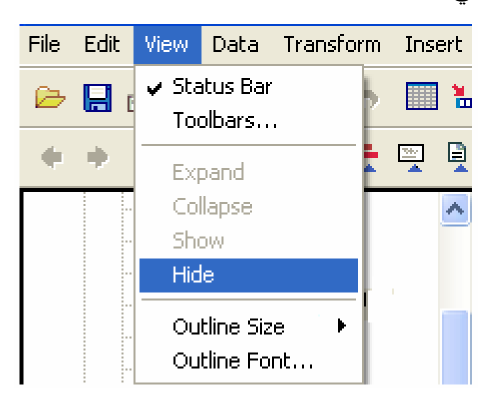
Model	Variables Entered	Variables Removed	Method
1	X[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

الشكل (12- 8)

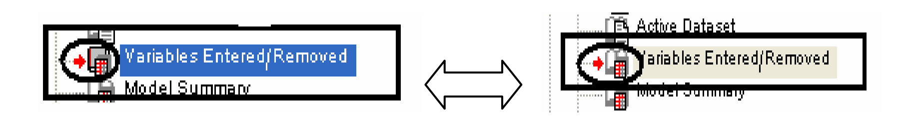
⌖ من القائمة View انقر الأمر Hide سيختفي الجزء التالي بمجرد النقر عليه كما في الشكل الاتي :



الشكل (12- 9)

⌖ لإظهار الجزء الذي اختفى انقر على الأمر Show من القائمة View ستظهر المحتويات مرة أخرى ونلاحظ أننا عند الإخفاء في أول مرة تكون Show غير منشطة وعندما نريد إظهار ما أردنا إخفاءه تكون Hide هي الغير نشطة بمعنى غير متاحة الاستخدام.

⌖ ملاحظه : يمكن الإخفاء والإظهار باستخدام القوائم Menu System ويتم ذلك بالنقر المزدوج على أيقونة الكتاب المجاورة للعنوان في الفهرس كما في الشكل الآتي فنلاحظ أنه عندما يكون الجدول متاح أو ظاهر فإن الكتاب يكون مفتوح وعندما يكون الجدول أو الشكل غير متاح فبنقر نقرا مزدوجا يصبح الكتاب مغلق والنقر عليه مرة أخرى يصبح مفتوح ومن ثم إظهار الشكل أو الجدول و العكس صحيح:



- **تحريك نتائج إلى مكان آخر:**

في صفحة النتائج يمكن تحريك Move أي محتوى من محتويات النتائج من مكانه إلى مكان آخر.

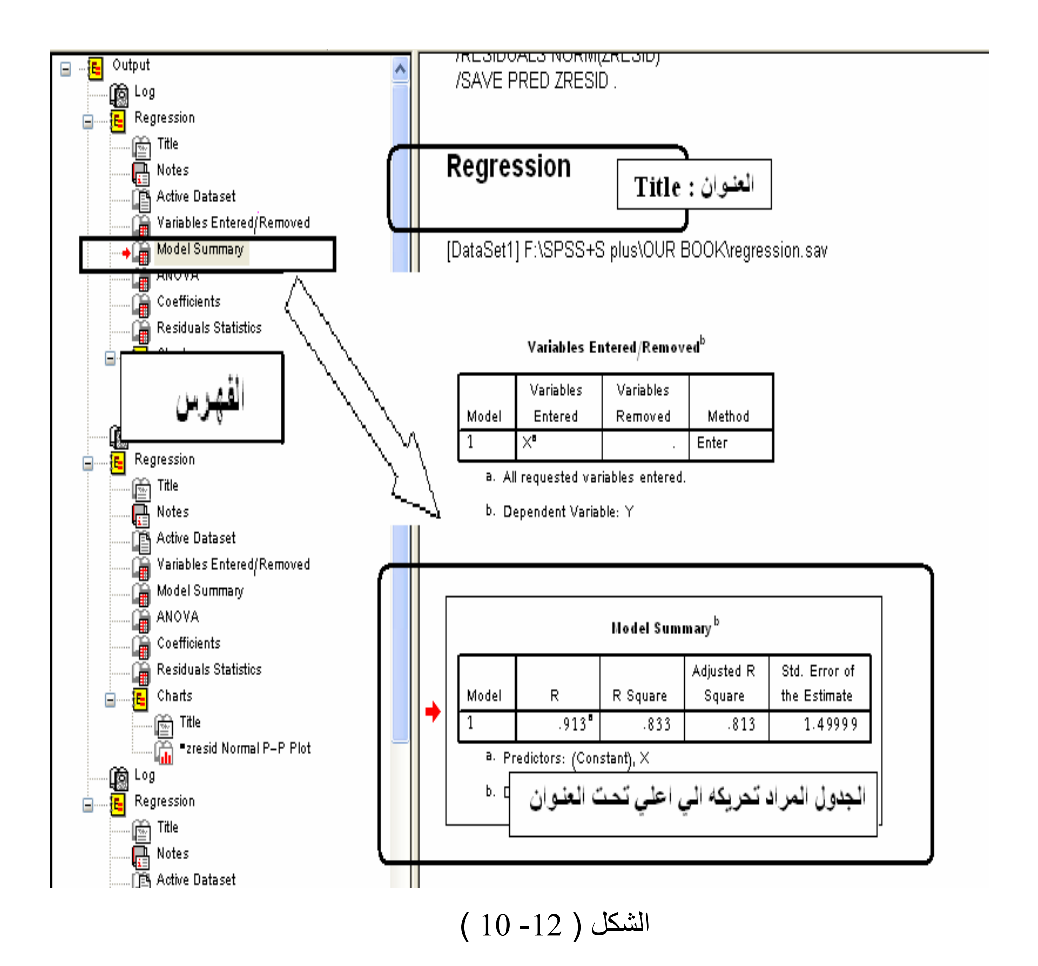
نفرض أننا نريد تحريك الجدول الذي بعنوان model summary إلى اعلى أسفل العنوان regression في المثال (11-1) فإننا نقوم بإتباع الخطوات الآتية:

⌖ انقر على الجدول المراد تحريكه في شاشة النتائج ونلاحظ أنه سيظهر السهم الأحمر والإطار الأسود وبذلك يكون الجدول نشطا.

⌖ سيضاء عنوان الجدول وهو المراد نقله في الفهرس يسارا.

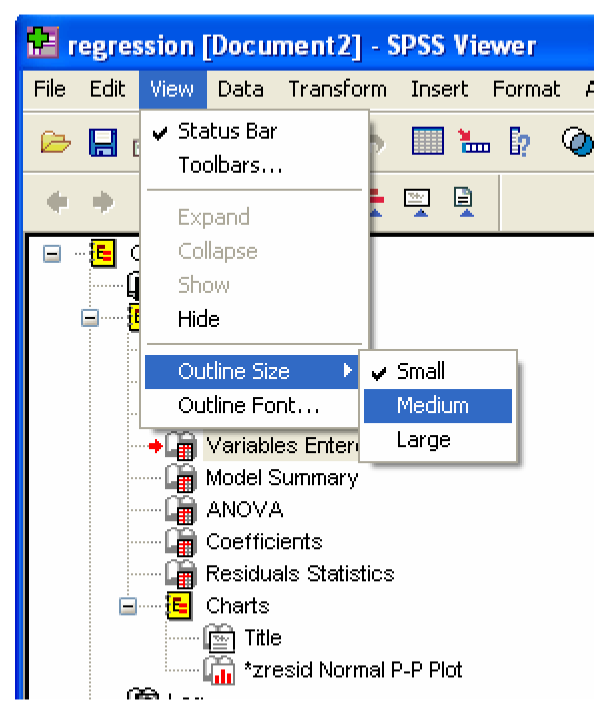
⌖ في الفهرس يسار الشاشة وباستخدام الزر الأيسر للفارة انقر على model summary واستمر في النقر مع تحريك الاسم في الفهرس إلى المكان المراد نقله إليه ونرفض أننا نريد نقله إلى أسفل العنوان Regression مباشراً .

⌖ عند الوقوف إلى المكان الذي نريد نترك الماوس فيتحرك الجدول إلى المكان الذي تم تحديده انظر الشكل (10-12).



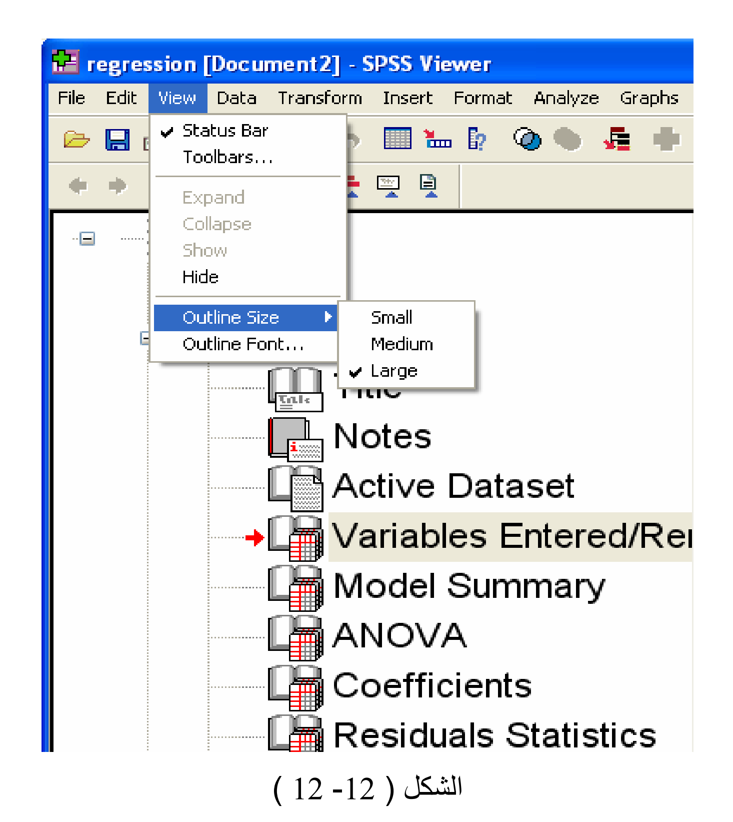
الشكل (12- 10)

- تغيير حجم وشكل الخط في الفهرس Change The Size and Font :
  - يمكن للمستخدم تغيير نوع الخط وحجمه في الفهرس بسهولة بالخطوات التالية:
  - من القائمة View انقر على الأمر Outline Size.
  - فيظهر قائمة يتم الاختيار منها كالآتي:



الشكل (12- 11)

⌖ فنجد من الشكل (12-11) إننا لدينا حجم صغير للخط في الفهرس وعرفنا ذلك من وجود علامة √ small وأن الحجم فعلا في الفهرس صغير وهذا واضـح مـن حجـم الخط وإذا أردنا أن نغيره إلى كبير فإننا ننقر على Large في الشكل (12-11) ومن ثم يصبح حجم الخط كبير كما في الشكل الآتي:

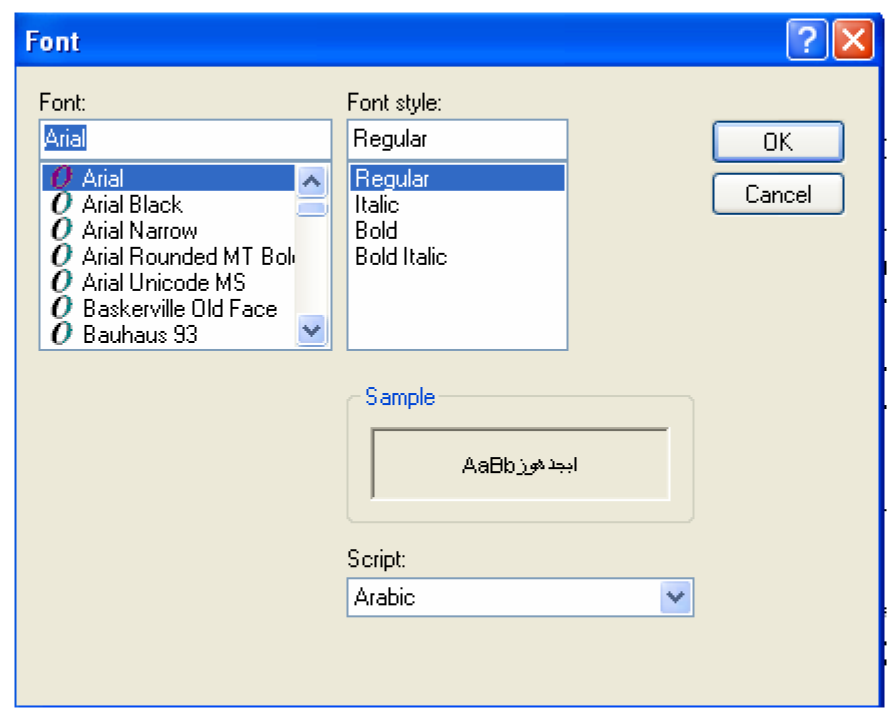


الشكل (12 -12)

كما أننا إذا أردنا أن نغير حجم الخط إلى متوسط فإننا ننقر على الأمر Medium سيتغير شكل الخط في الفهرس من الكبير إلى متوسط.

ونجد أيضا أننا يمكننا من تغيير نوع الخط في الفهرس كما يلي:

- من القائمة View انقر الأمر Font Outline سيفتح صندوق به العديد من الأنواع انقر على النوع الذي ترغبه من الخطوط انظر الشكل (12 -13):



الشكل (12- 13)

⌖ وبعد الانتهاء من الاختيار يتم النقر على الأمر Ok سيتغير نوع الخط بمجرد التنفيذ.

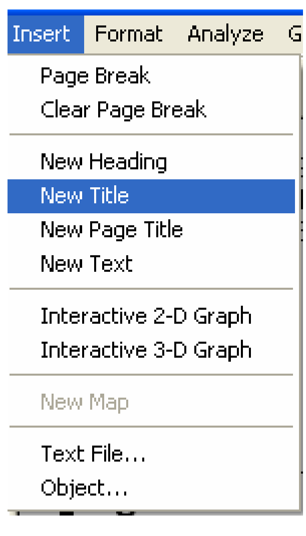
(4-12) أضافة عنوان جديد Add New Title:

مثال (1-12):

في المثال (2-11) نفرض أننا نريد إضافة عنوان أسفل العنوان Regression بعنوان جديد يسمى" نموذج الانحدار المتعدد" فإننا نتبع الخطوات الآتية:

انقر على الشكل المراد أضافة عنوان أسفله حتى يضاء وهنا في المثال وهو العنوان Regression.

من القائمة Insert انقر على الأمر New Title نتيجة تنفيذ الأمر سينشأ البرنامج مستطيلا أسفل الشكل كما في الشكل الآتي:



الشكل (12- 14)

- أنقر نقرا مزدوجا على المستطيل الجديد.
- اكتب العنوان المراد إدخاله للصفحة في المستطيل وهو " الانحدار المتعدد" فيظهر الشكل الآتي:

Regression

الانحدار المتعدد

[DataSet1] F:\SPSS+S plus\OUR BOOK\regression2.sav

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	WorkHours[a]	.	Enter

a. All requested variables entered.

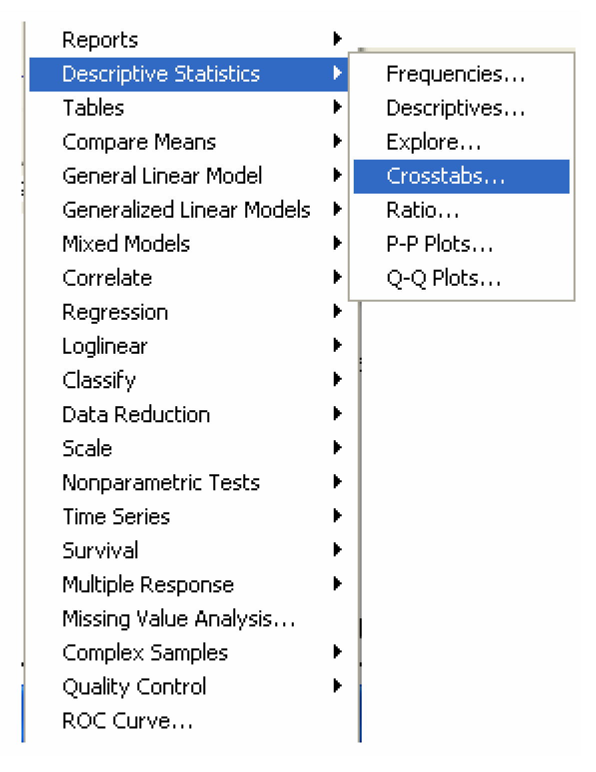
b. Dependent Variable: SpareParts

الشكل (12- 15)

⌖ يمكن إزاله العنوان القديم وتحريك العنوان الجديد في مكان العنوان القديم.

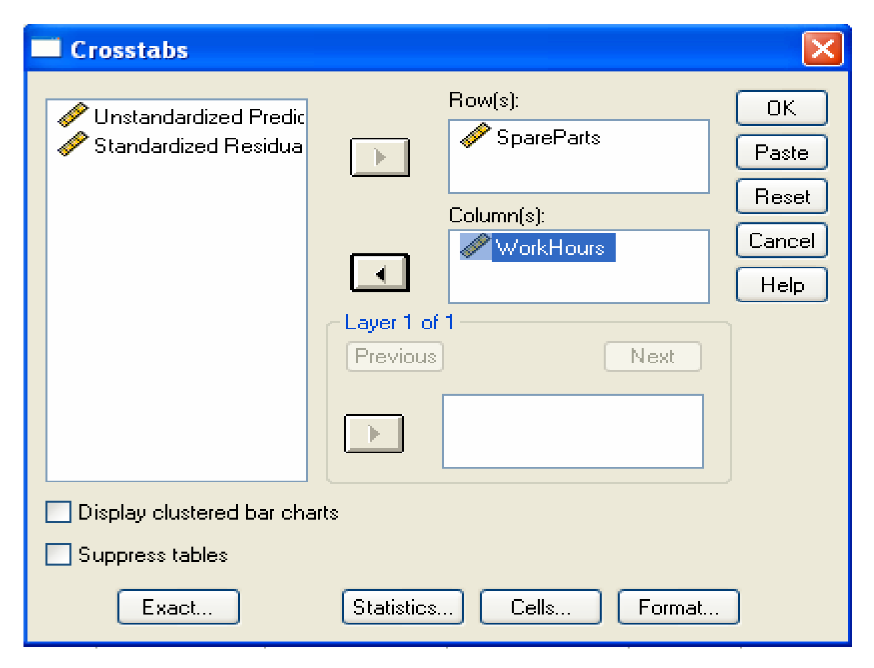
(5-12) محرر الجداول Pivot Tables:

عند فتح صفحة البيانات المثال (2-11) ثم من القائمة Analyze ثم Descriptive Statistics والأمر Cross Tab من الشكل الآتي:



الشكل (12- 16)

بعد ذلك يظهر مربع الحوار الآتي ونقوم بإدخال المتغير الذي نريد أن يظهر في الصف ونفرض أنه SpareParts ونفرض أن المتغير المراد إدخاله في الأعمدة هو WorkHours:



الشكل (12- 17)

فنحصل على جدول التوزيع المزدوج نتيجة تنفيذ الأمر الآتي:

- جدول به المتاح و المفقود من البيانات.

جدول (12- 1)

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SpareParts * WorkHo	10	100.0%	0	.0%	10	100.0%

ب- جدول التوزيع الثنائي المطلوب.

SpareParts * WorkHours Crosstabulation

Count

		WorkHours										
		50.00	69.00	73.00	87.00	108.00	128.00	132.00	135.00	148.00	170.00	Total
SpareParts	20.00	1	0	0	0	0	0	0	0	0	0	1
	30.00	0	1	1	0	0	0	0	0	0	0	2
	40.00	0	0	0	1	0	0	0	0	0	0	1
	50.00	0	0	0	0	1	0	0	0	0	0	1
	60.00	0	0	0	0	0	1	1	1	0	0	3
	70.00	0	0	0	0	0	0	0	0	1	0	1
	80.00	0	0	0	0	0	0	0	0	0	1	1
Total		1	1	1	1	1	1	1	1	1	1	10

لتعديل جدول معين أو بمعنى آخر لإعداد الجدول ليستجيب إلى أية تعديلات يطلبها المستخدم نمر بالخطوات التالية:

1-ننقر نقرا مزدوجا على عنوان الجدول المراد التعديل فيه وهو (SpareParts * WorkHours Crosstabulation) .

هذا النقر سيجعل الجدول نشطا (انظر الشكل 12-18):

SpareParts * WorkHours Crosstabulation

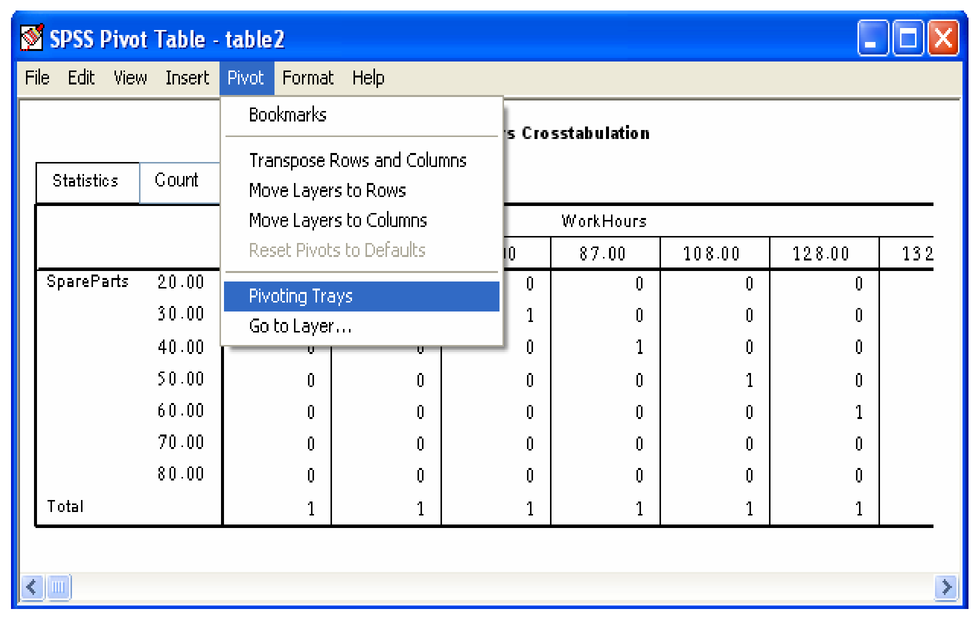
Count

		WorkHours										
		50.00	69.00	73.00	87.00	108.00	128.00	132.00	135.00	148.00	170.00	Total
SpareParts	20.00	1	0	0	0	0	0	0	0	0	0	1
	30.00	0	1	1	0	0	0	0	0	0	0	2
	40.00	0	0	0	1	0	0	0	0	0	0	1
	50.00	0	0	0	0	1	0	0	0	0	0	1
	60.00	0	0	0	0	0	1	1	1	0	0	3
	70.00	0	0	0	0	0	0	0	0	1	0	1
	80.00	0	0	0	0	0	0	0	0	0	1	1
Total		1	1	1	1	1	1	1	1	1	1	10

الشكل (12- 18)

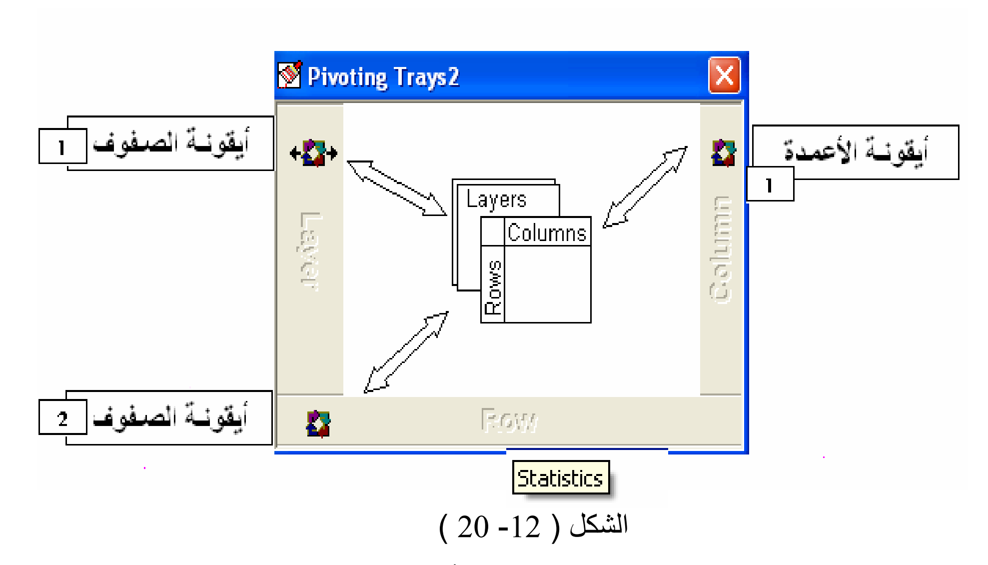
أ-لتبديل الصفوف بالأعمده To Transpose Rows And Columns:

بعد تنشيط الجدول ومن القائمة Pivot انقر الأمر Pivot Trays كم يلي :



الشكل (12- 19)

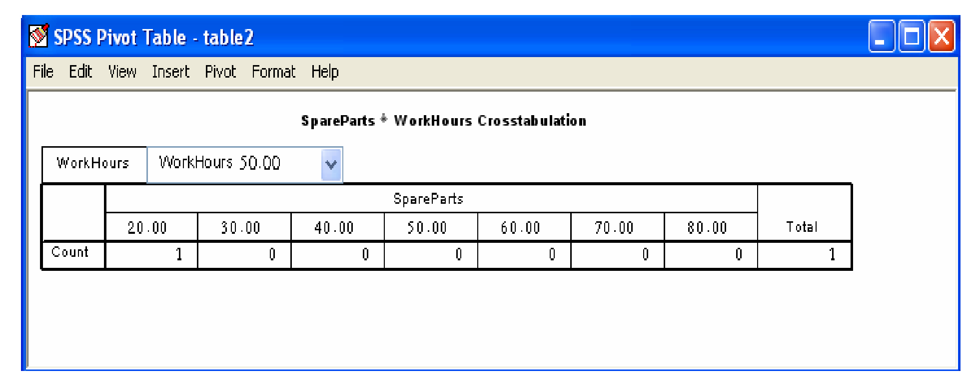
لفتح الصندوق الخاص بذلك وهو يسمى Pivot Trays1 (أنظر الشكل 12-20):



الشكل (12- 20)

يوجد بالصندوق الحواري Pivot Trays 1 ثلاثة أيقونات اثنان على يسار الصندوق بجوار كلمة Row وواحدة بجوار كلمة Column الأيقونة التي على اليمين تمثل المتغير (SpareParts) والأيقونات الأفقية أحدهما تمثل WorkHours.

2- بالفأرة اسحب الأيقونة الرأسية بجوار كلمة Column إلى الصف الأفقي واسحب أيقونه الصفوف الأفقية بجوار Row إلى الأعمدة.



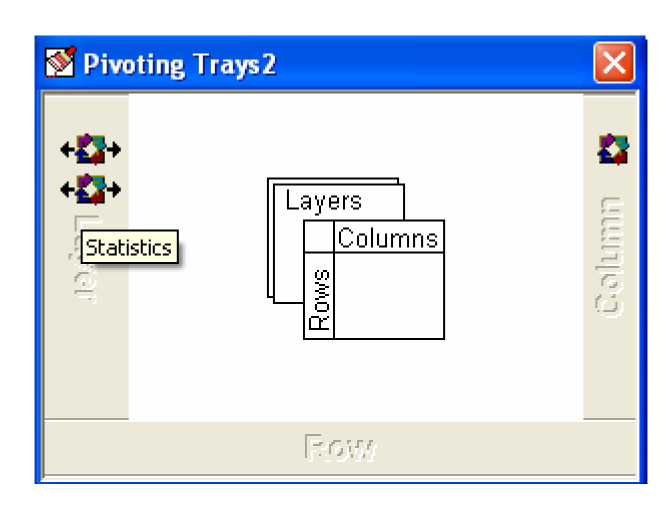
	SpareParts							Total
	20.00	30.00	40.00	50.00	60.00	70.00	80.00	
Count	1	0	0	0	0	0	0	1

الشكل (12- 21)

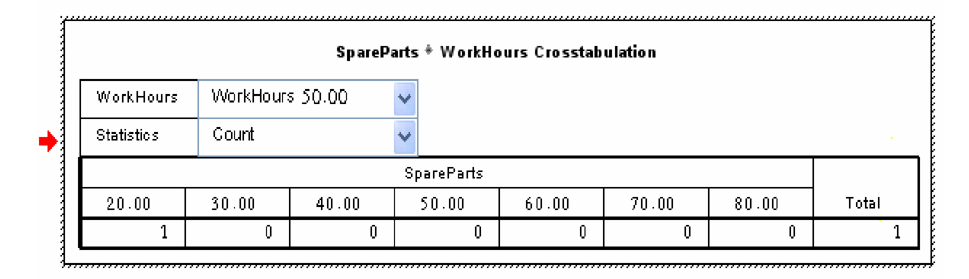
ب- لتغيير ترتيب العرض في الجدول المحور To Change The Display Order In The Pivot Table :

تتم هذه العملية كالتالي :

- تأكد أن الجدول المحور تم تنشيطه وظهر على الشاشة الصندوق Pivoting Tray1.
- في الصندوق الحواري Pivoting Tray1 انقر الأيقونة Statistics واسحبها إلى يمين شريط الصف ستلاحظ أن ترتيب العناصر داخل الجدول تغير وأصبح مفهوما كالآتي (انظر الشكل12-22):



الشكل (12- 22)



SpareParts * WorkHours Crosstabulation

WorkHours	WorkHours 50.00
Statistics	Count

SpareParts							Total
20.00	30.00	40.00	50.00	60.00	70.00	80.00	
1	0	0	0	0	0	0	1

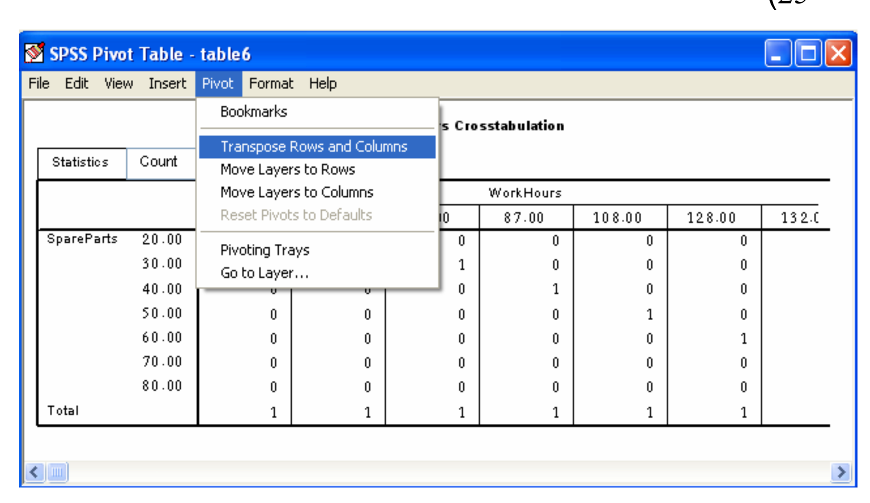
الشكل (12- 23)

في الشكل (12-23) نلاحظ أن الجدول تكون تحت فئة واحدة وهي WorkHours 50.00 وبالنقر على السهم يظهر الفئات جميعا وباختيارها يظهر SpareParts بفئة WorkHours الجديدة.

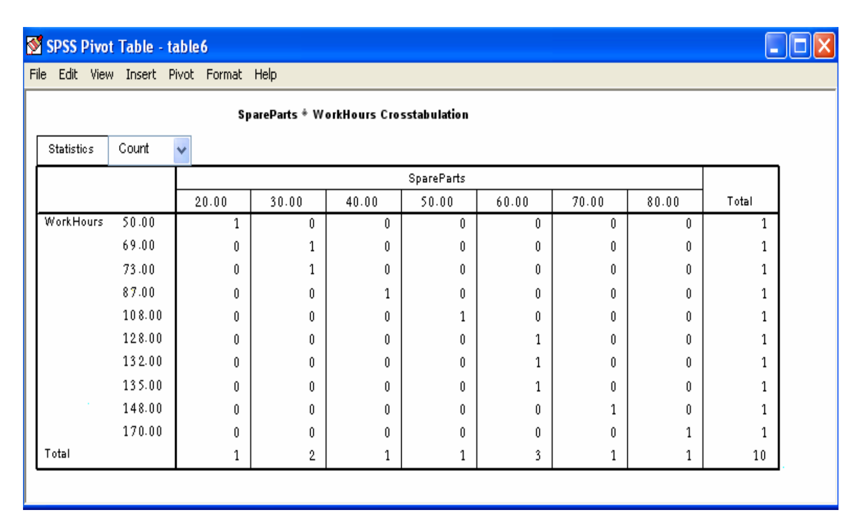
جـ- لتحريك صف أو عمود To Move Rows & Column :

لتحريك صف أو عمود في الجدول تتم العملية كالتالي:

- نقوم بتنشيط شكل رقم (12-21) بنفس الخطوات السابقة أي بالنقر على الجدول المراد نقل أعمدة وتبديلها .
- من Pivot نختار Transpose Rows and Colums كما في الشكل (12-24) ومن ثم يتم تحويل الصف عمود و العمود إلى صف كما في الشكل (12-25)



الشكل (12- 24)



SPSS Pivot Table - table6

File Edit View Insert Pivot Format Help

SpareParts * WorkHours Crosstabulation

Statistics: Count

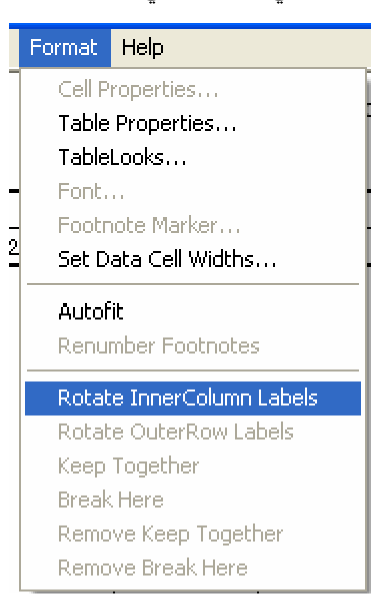
		SpareParts							
		20.00	30.00	40.00	50.00	60.00	70.00	80.00	Total
WorkHours	50.00	1	0	0	0	0	0	0	1
	69.00	0	1	0	0	0	0	0	1
	73.00	0	1	0	0	0	0	0	1
	87.00	0	0	1	0	0	0	0	1
	108.00	0	0	0	1	0	0	0	1
	128.00	0	0	0	0	1	0	0	1
	132.00	0	0	0	0	1	0	0	1
	135.00	0	0	0	0	1	0	0	1
	148.00	0	0	0	0	0	1	0	1
	170.00	0	0	0	0	0	0	1	1
Total		1	2	1	1	3	1	1	10

الشكل (12- 25)

د- لتدوير مميز صف أو عمود To Rotate Row Or Column Label :

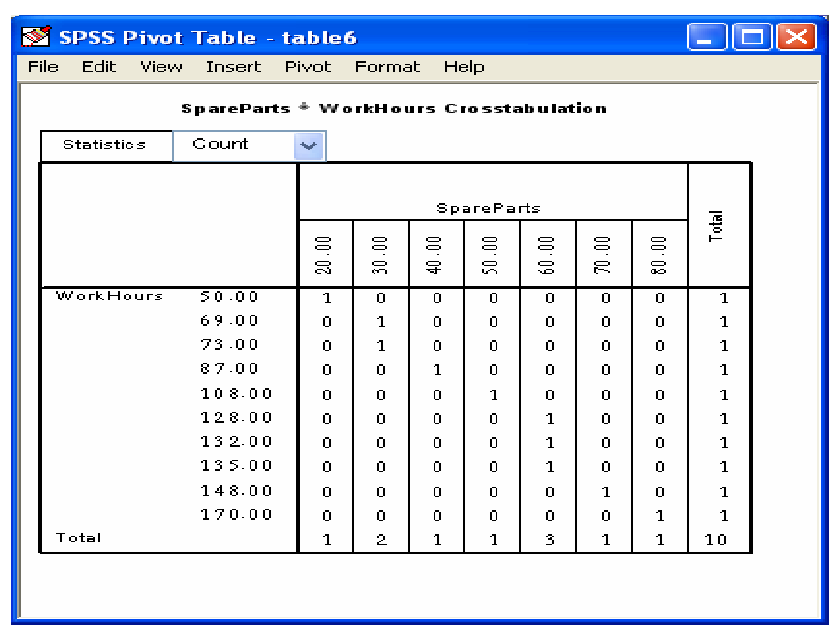
⌖ نشط الجدول المحور.

○ من القائمة Format انقر الأمر Rotate Inner- Column Labels كما في الشكل الآتي:



الشكل (12- 26)

⌖ ستلاحظ أن مميز الأعمدة تم تدويره و يمكن استخدام نفس الأمر للصفوف انظر الجدول التالي.



SPSS Pivot Table - table6

File Edit View Insert Pivot Format Help

SpareParts * WorkHours Crosstabulation

Statistics: Count

		SpareParts							Total
		20.00	30.00	40.00	50.00	60.00	70.00	80.00	
WorkHours	50.00	1	0	0	0	0	0	0	1
	69.00	0	1	0	0	0	0	0	1
	73.00	0	1	0	0	0	0	0	1
	87.00	0	0	1	0	0	0	0	1
	108.00	0	0	0	1	0	0	0	1
	128.00	0	0	0	0	1	0	0	1
	132.00	0	0	0	0	1	0	0	1
	135.00	0	0	0	0	1	0	0	1
	148.00	0	0	0	0	0	1	0	1
	170.00	0	0	0	0	0	0	1	1
Total		1	2	1	1	3	1	1	10

الشكل (12- 27)

هـ - تغيير مظهر الجدول Change The Appearance Of Tables:

يمكن تغيير مظهر الجدول بتطبيق النظر للجدول Table Look ولتطبيق النظر للجدول نجري الآتي:

⌖ نشط الجدول المحور بالطرق المعروفة.

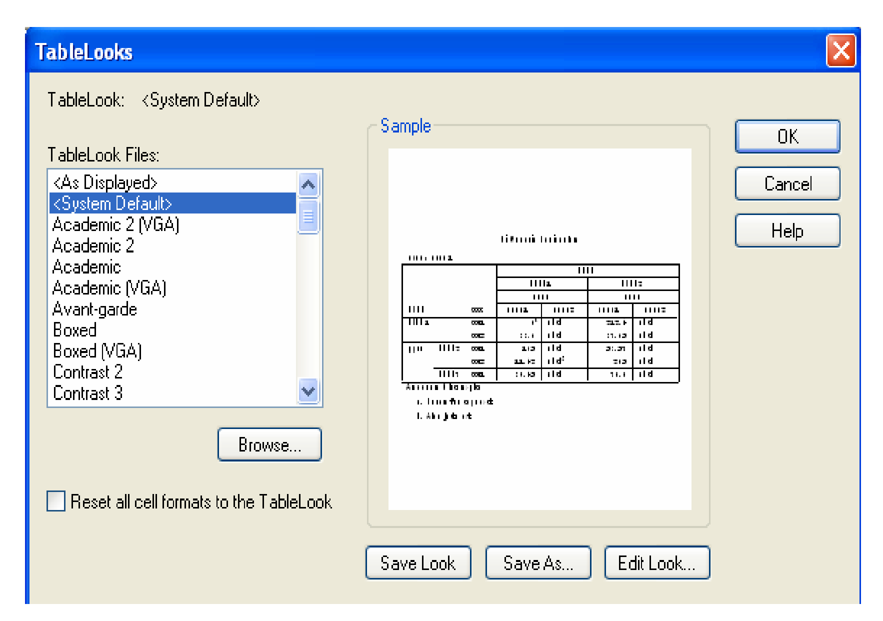
⌖ من القائمة Format انقر الأمر Table Look.

⌖ اختر المظهر الملائم و الذي ترغب فيه من قائمة الملفات.

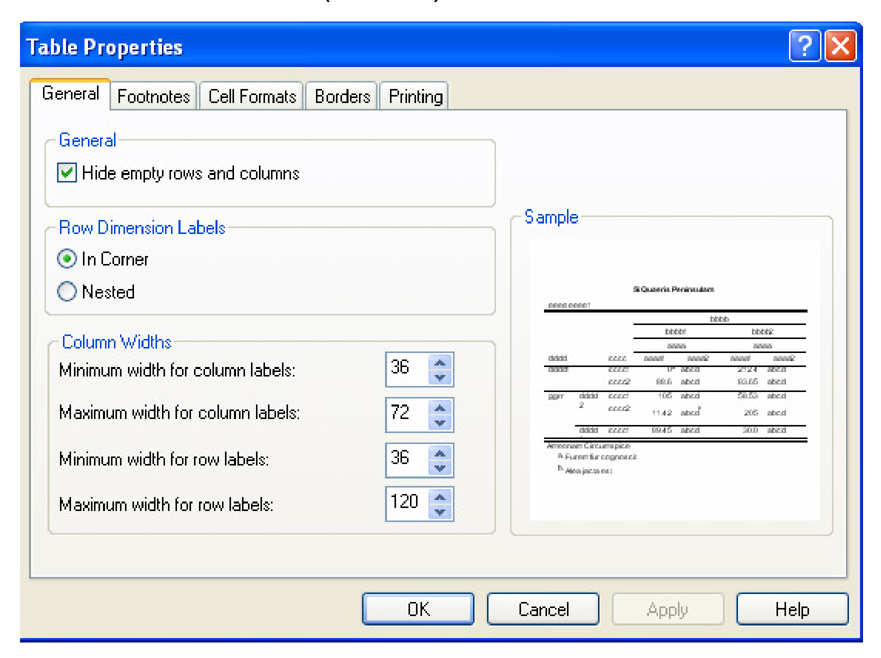
⌖ لتطبيق الشكل المختار على الجدول انقر الأمر Ok انظر الشكل (12-28).

⌖ ملاحظة:

⌖ يمكن تغيير خصائص الجدول خلال قائمة Format باستخدام الخيارات Table Properties ، ومن ثم تغير خصائص Cell والحدود والتظليل و أشياء أخرى كثيرة مرتبطة بالتنسيق (انظر الشكل 12-29).



الشكل (12- 28)



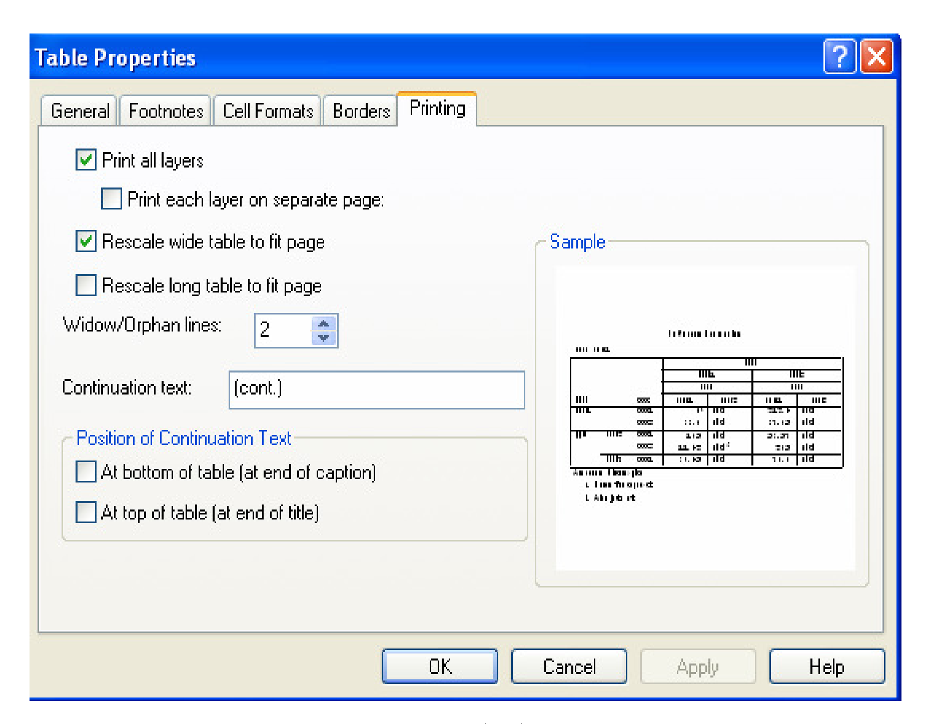
الشكل (12- 29)

و- طباعة الجداول المحور Printing Pivot Table:

لطباعة تقسيم مخفي من الجدول وتحجيم الجدول ليلائم حجم الصفحة نمر بالخطوات التالية :

نشط الجدول بالطرق المعروفة.

- من القائمة Format انقر الأمر Table Properties لفتح الصندوق Table Properties الخاص بالعملية الموجود بالصفحة التالية.
- انقر على الأمر Printing فيظهر مربع حوار.
- بالأفرة اختار الصندوق Print All Layers ستلاحظ أن هنا صندوقاً يطلب هل الطباعة تتم كل تقسيم في صفحة مستقلة.
- لتحجيم الجدول انقر Rescale Wide Table To Fit Page .
- انقر على الأمر Ok للتنفيذ انظر الشكل (30-12).



الشكل (30-12)

(6-12) تعديل وتنقيح الأشكال البيانية Modifying And Enhancing Charts Presentation:

قد يرغب المستخدم في إجراء تعديلات على الشكل البياني بتغيير الألوان أو نوع الخط أو إضافة معلومات هامه لبحثها مثل:

1 - إضافة عنوان آخر غير العنوان الغيابي.

2 - تعديل الأعمدة البيانية لتكون ثلاثية الأبعاد 3-D أو تغيير المسافات بين الأعمدة و خلافه.

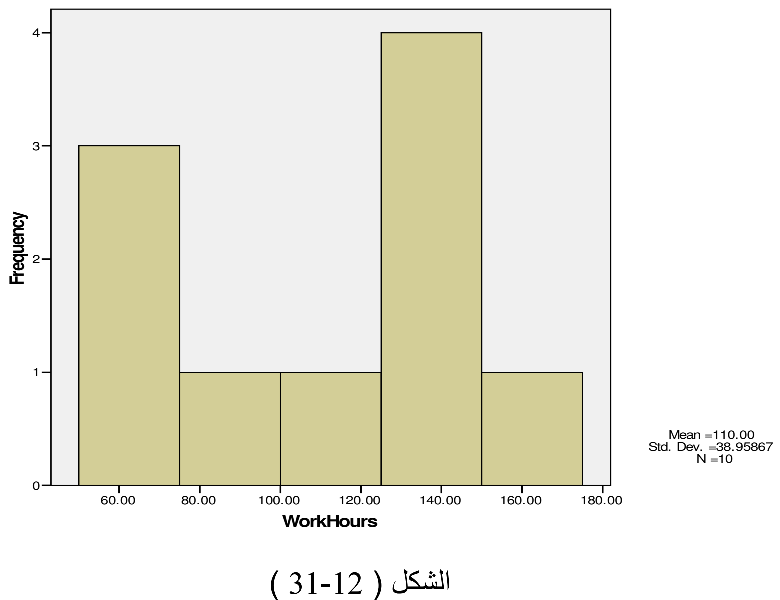
3 - حذف أو إضافة عناوين جديدة أو التغير بطريقة تبرز خصائص الشكل البياني بطريقة أحسن.

إضافة مقاييس إحصائية.

نفرض أن لدينا العمود البياني من المثال (2-11) وهو عمود ساعات العمل :

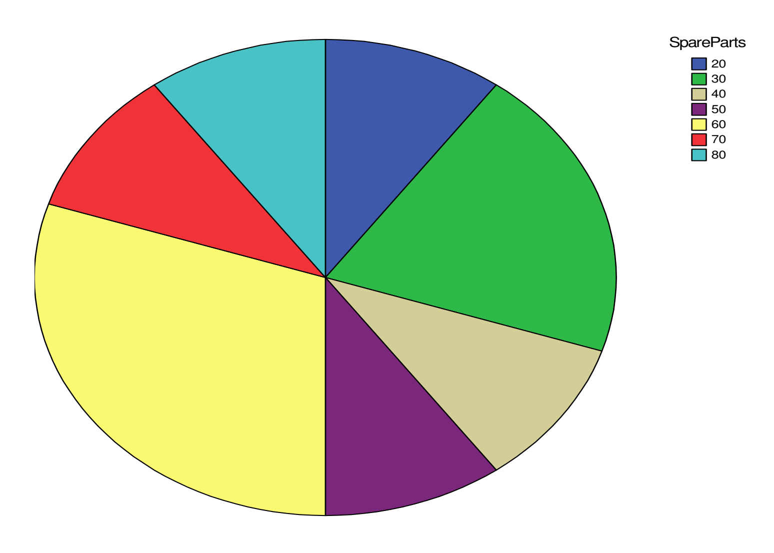
WorkHours
73.00
50.00
128.00
170.00
87.00
108.00
135.00
69.00
148.00
132.00

ونفترض أننا رسمنا المدرج التكراري لساعات العمل كما يلي ونلاحظ أننا رسمناه بنفس طريقة رسم الأشكال انظر الفصل الثاني:



الشكل (12-31)

وتم الحصول أيضا على الدائرة التي تمثل المتغير (SpareParts) ونلاحظ أننا غيرنا نوع المتغير إلى ترتيبي ordinal لإنشاء دائرة منها و هي كالآتي (انظر شكل 12-32):

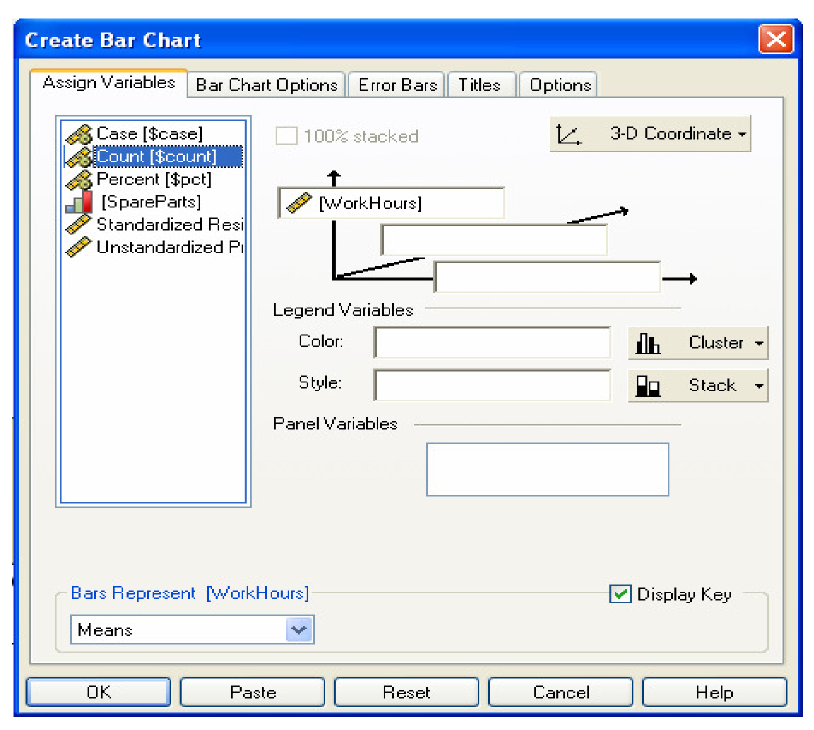


الشكل (12-32)

أ- تغيير أبعاد الأعمدة إلى البعد الثلاثي 3-D:

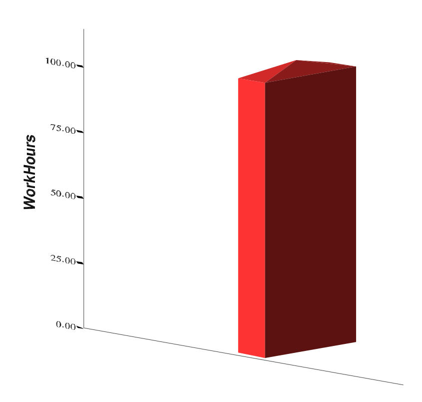
يمكن إضافة بعد ثالث للرسم من خلال محرر الشكل كالآتي:

- يتم الحصول على الشكل البياني المراد تعديله.
- انقر على الشكل المراد تعديله.
- من القائمة Graphs انقر على Interaction ثم Bar .
- انقر على الأمر Effect 3-D ستلاحظ أن أبعاد الأعمدة أصبحت ثلاثية انظر الصندوق الفرعي التالي (انظر الشكل 12- 33) وكذلك الأعمدة بعد التغيير انظر الشكل(12-34):



الشكل (12- 33)

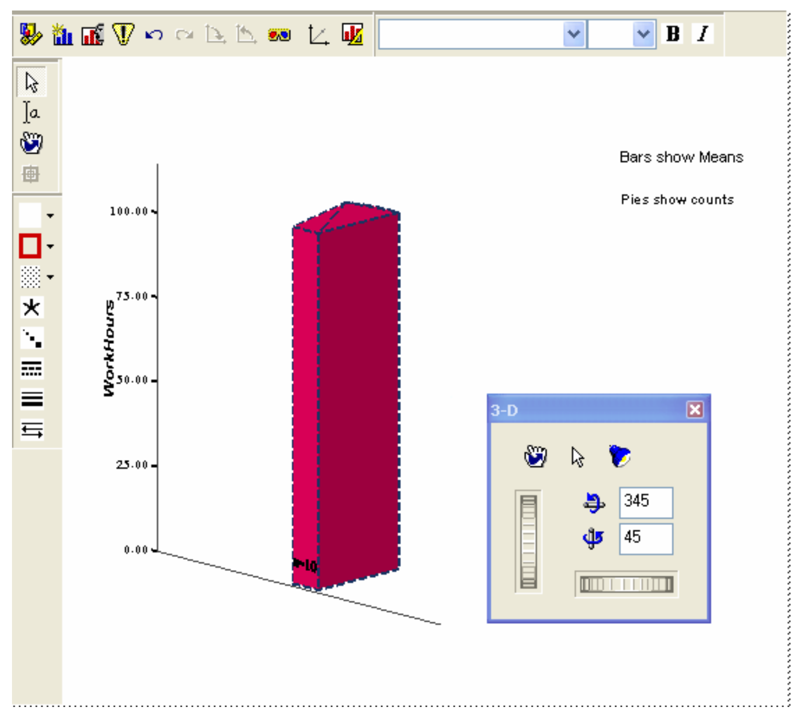
Bars show Means



الشكل (12- 34)

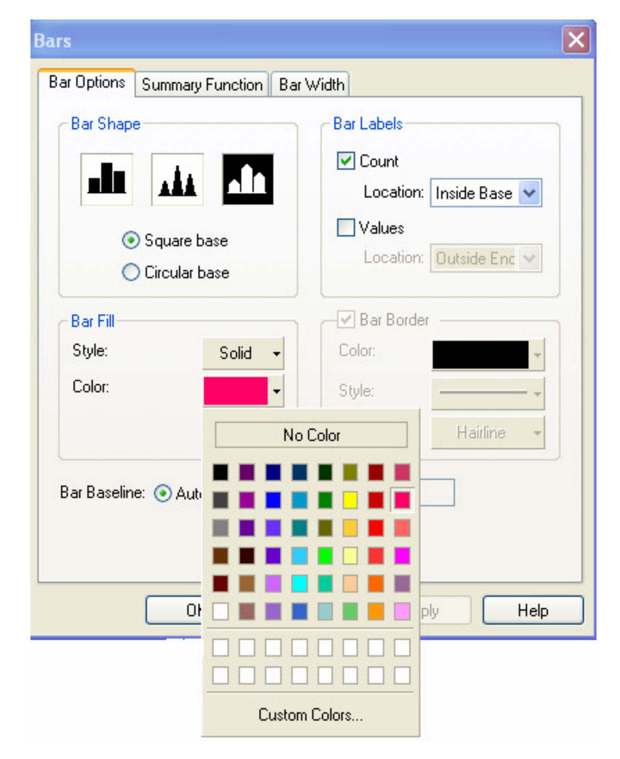
ب – تعديل خصائص الشكل Chart Attributes:

- انقر نقرا مزدوجا على أي مكان على الشكل سيفتح الصندوق الخاص بمحرر الشكل.
- عند النقر المزدوج على أي عنوان في الشكل سيفتح الصندوق الذي يسمح لك بضبط النص وإجراء ما تريده من الشكل يمكن للمستخدم تغيير العنوان (يمينا أو يسارا أو في الوسط) وكذلك تغيير المميز (أنظر الشكل 12- 35).

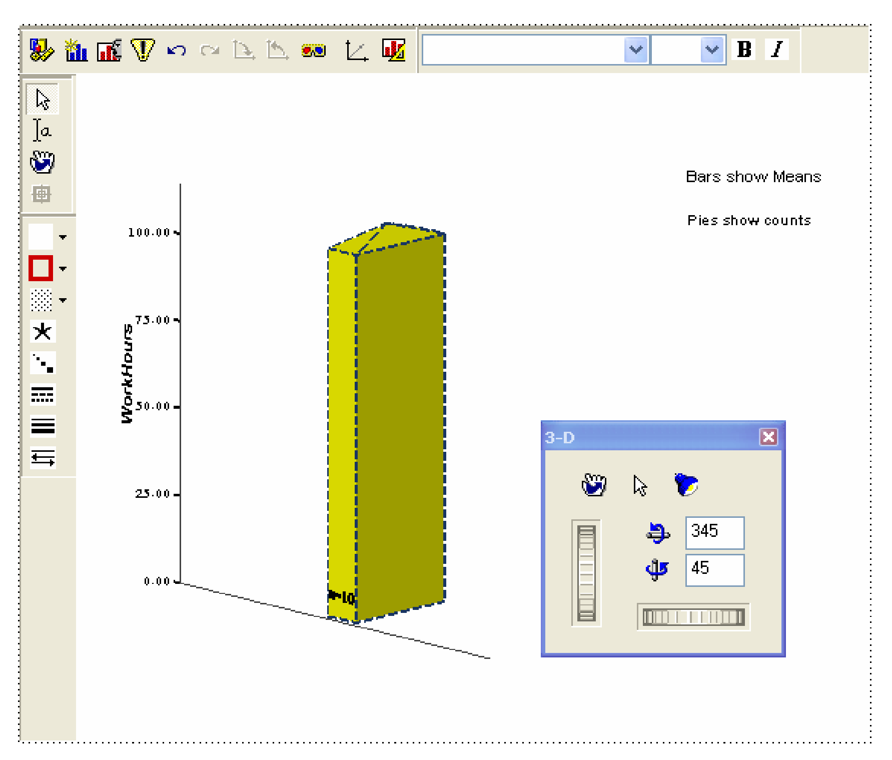


الشكل (12- 35)

⌖ وننقر على أي خط أو رسمة في الشكل (12-35) مرتين متتالين يظهر مربع حوار خاص بخصائصه من تغير عرضه وتغير لونه وهكذا فإذا افترضنا أننا نريد لون تغير الشكل من اللون الأحمر إلى الأصفر فإننا ننقر على الشكل الذي باللون الأحمر مرتين متتالين فيظهر الشكل الآتي:


Bars
Bar Options
Summary Function
Bar Width
Bar Shape
Square base
Circular base
Bar Labels
Count
Location:
Inside Base
Values
Location:
Outside End
Bar Fill
Style:
Solid
Color:
Bar Border
Color:
Style:
Hairline
No Color
Bar Baseline:
Custom Colors...
Help

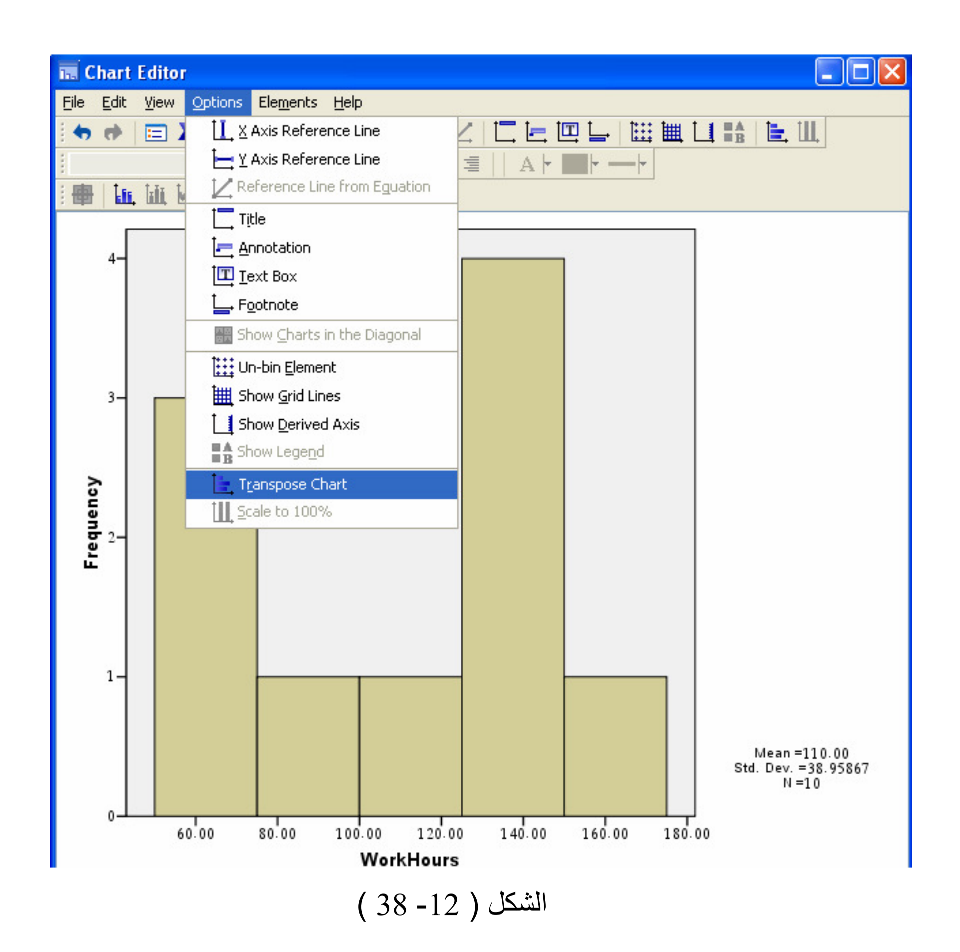
الشكل (12- 36)



الشكل (12- 37)

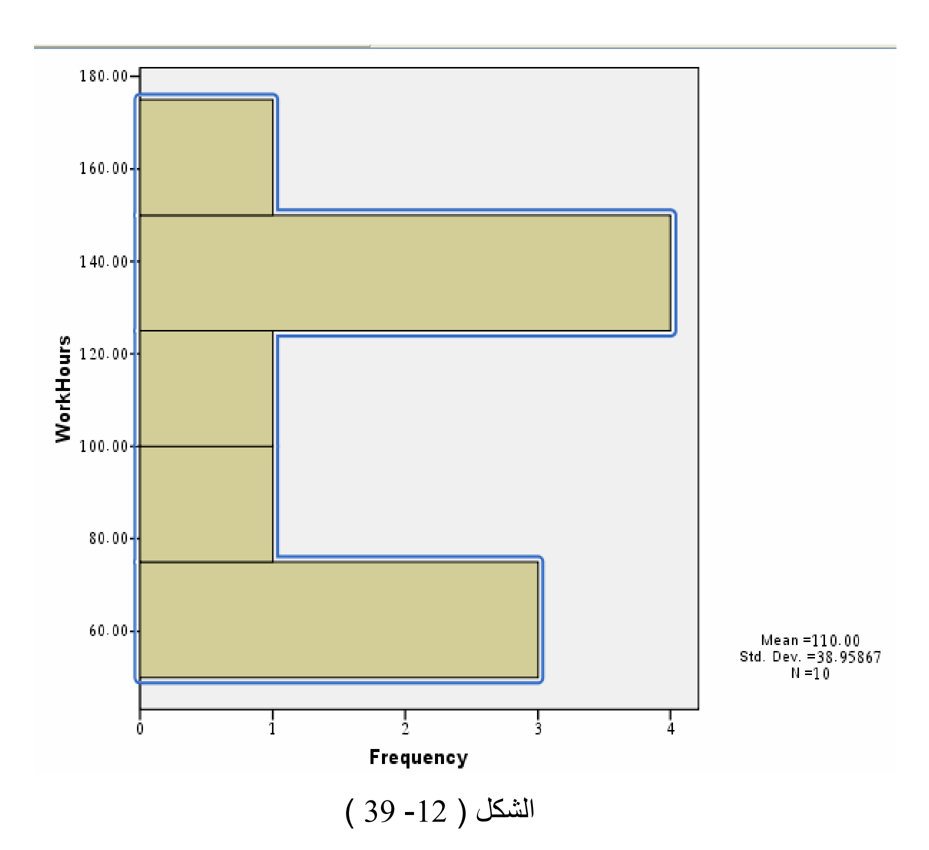
ملاحظة:

- ✓ تذكر دائما عند تغيير مواصفات الشكل أن تنقر على الأمر ok لتنفيذ الأمر أو cancel لإغلاق الصندوق.
- ✓ د- لتبادل المتغيرات على الشكل To Transpose Variables In A Chart:
- ✓ ننقر على الشكل مرتين متتاليين و فيظهر مربع حوار منه يتم اختيار option:



الشكل (12- 38)

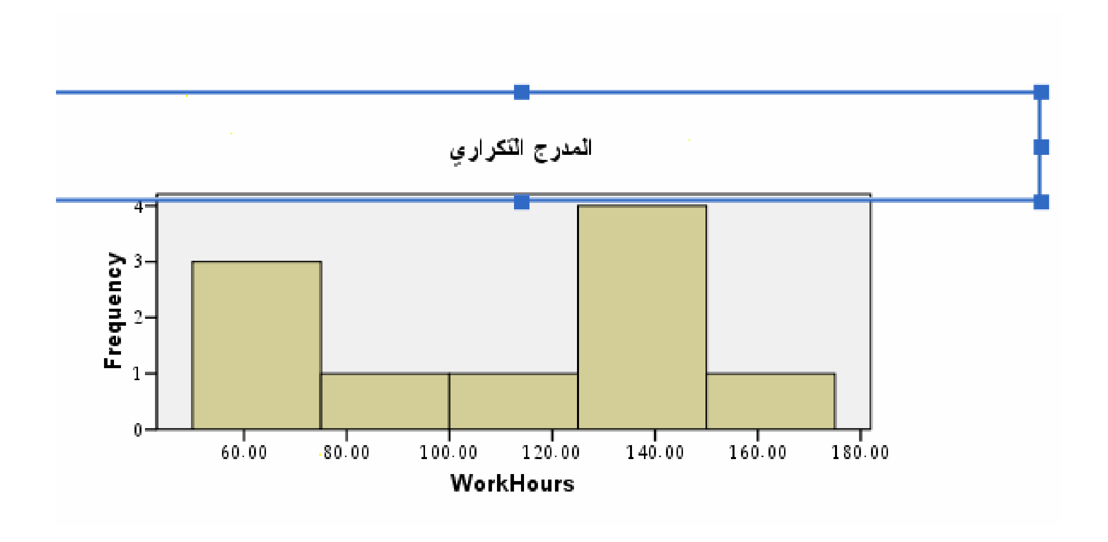
⌖ من القائمة Options انقر على الأمر Transpose فيصبح المتغير الموجود على المحور الأفقي على المحور الرأسي والعكس (انظر الشكل 12- 39):



الشكل (12- 39)

هـ - لإدراج عنوان جديد للشكل To Insert A New Chart Title:

- من القائمة Options انقر على Title فيتم تنشيط عنوان أعلي الشكل لكي يتيح لنا كتابة عنوان به.
- يمكن عن طريق لوحة المفاتيح إضافة العنوان الجديد كما في الشكل الآتي إذا أضفنا عنوان على الشكل (12-19) بعنوان" المدرج التكراري: فيصبح الشكل (12-31) كالآتي:


المدرج التكراري
Frequency
0
1
2
3
4
60.00
80.00
100.00
120.00
140.00
160.00
180.00
WorkHours

المراجع

المراجع العربية:

- الهلباوي، عبد الله توفيق، الإحصاء التطبيقي ، مكتبة عين شمس، 2000.
- بشير، سعد زغلول، دليلك إلى البرنامج الإحصائي SPSS،، الإصدار العاشر، الجهاز المركزي للإحصاء جمهورية العراق.
- عاشور،سمير كامل وسالم،سامية أبو الفتوح ، العرض والتحليل الإحصائي بإستخدام SPSSWIN 2003.
- هندي،محمود محمد إبراهيم وسلمان،خلف سلمان سلطان،مفاهيم لطرق التحليل الإحصائي ،مكتبة الرشد،الرياض،1425.
- سلطان، عبد الله علي حسن و حسين، علي الهفوفي، اساسيات العرض والتحليل الإحصائي باستخدام SPSS-WIN (الجزء الأول). مشروع تخرج بجامعه الملك سعود، كليه العلوم، قسم الإحصاء وبحوث العمليات، 1427 هـ.
- ثامر، محمد عباس منشي ودهام، ناصر دهام الدهام، اساسيات العرض والتحليل الإحصائي باستخدام SPSS-WIN (الجزء الثاني). مشروع تخرج بجامعه الملك سعود، كليه العلوم، قسم الإحصاء وبحوث العمليات، 1428 هـ.

المراجع الأجنبية:

- Hollander، M. & Wolf، D. A. (1998). Nonparametric Statistical Methods، Second Edition، Wiley، New York.
- Gupta، V. (1999). SPSS for Beginners. VJBooks Inc.

الفهرس